Nanotechnology

Self-assembling DNA recognizes patterns

Andrew Phillips

Self-assembling DNA can process information, but the computations have been limited to digital algorithms. A self-assembling DNA system has now been designed to perform complex pattern recognition. **See p.500**

How do biological systems process information? Such computations occur at multiple scales, involving molecules, cells, tissues, organs and beyond. For example, computational circuits are embedded in molecular pathways[1] and formed by the neurons of the brain[2], and the immune system computes immune responses by processing a vast array of molecular inputs[3]. Understanding such information processing could unlock strategies for treating disease and identify fundamental engineering principles of biology.

In human-engineered systems, computations typically take place in dedicated processing units. By contrast, the computations involved in biological processes such as protein synthesis, molecular self-assembly and cell navigation are often embedded directly in biophysical processes. Understanding such biophysically embedded computations remains a key challenge. On page 500, Evans *et al.*[4] report an investigation of biophysical computations in which a system of self-assembling DNA tiles was used to recognize patterns.

The authors' work builds on the field of DNA nanotechnology[5], which uses DNA to construct sophisticated nanostructures for a broad range of applications. It takes advantage of the fact that single strands of DNA can be designed to self-assemble in solution by binding to DNA molecules with complementary sequences. One strategy is to assemble structures from DNA building blocks in the form of square tiles, which are designed so that tile edges with complementary sequences bind to each other. Such self-assembling tile systems have been engineered to implement a broad range of digital algorithms[6], but other sophisticated computations, such as pattern recognition, have not been explored with these systems.

To implement complex pattern-recognition computations, Evans *et al.* designed a set of square DNA tiles to self-assemble into three distinct shapes: the letters H, A and M (Fig. 1a). Some of the tiles were used in the assembly of only one of the shapes; others were used in two or three shapes. The tile set was designed so that the three shapes formed in similar quantities when the different tiles were present in solution in equal concentrations. By contrast, different concentrations of each type of tile resulted in the formation of one dominant shape, which depended on the concentrations.

The full tile system was designed computationally, in a process that started with each of the three shapes composed entirely of tiles unique to that shape, each tile having a distinct set of edges. A merging algorithm then progressively identified tiles that could be used in more than one shape. Each step of the algorithm randomly picked two tiles from two different shapes and tried to find one tile that could replace them both, altering the neighbouring edges accordingly. A replacement was accepted only if the resulting system satisfied certain conditions, to prevent the system from getting 'stuck' in incorrect assemblies. The conditions assume that tiles bound on a single edge readily fall off the self-assembling shape, whereas those bound on multiple edges are stable. When the conditions are met, incorrect tiles can bind to the system only on a single edge; if an incorrect tile does bind, a subsequent incorrect tile can also bind only on a single edge. After multiple iterations of the algorithm, an optimized system was obtained, consisting of 917 tiles – 371 of which are used in more than one shape – with 698 distinct edges.

The next challenge was to identify groups of the tiles that co-locate in only one of the three assembled shapes to form the core of

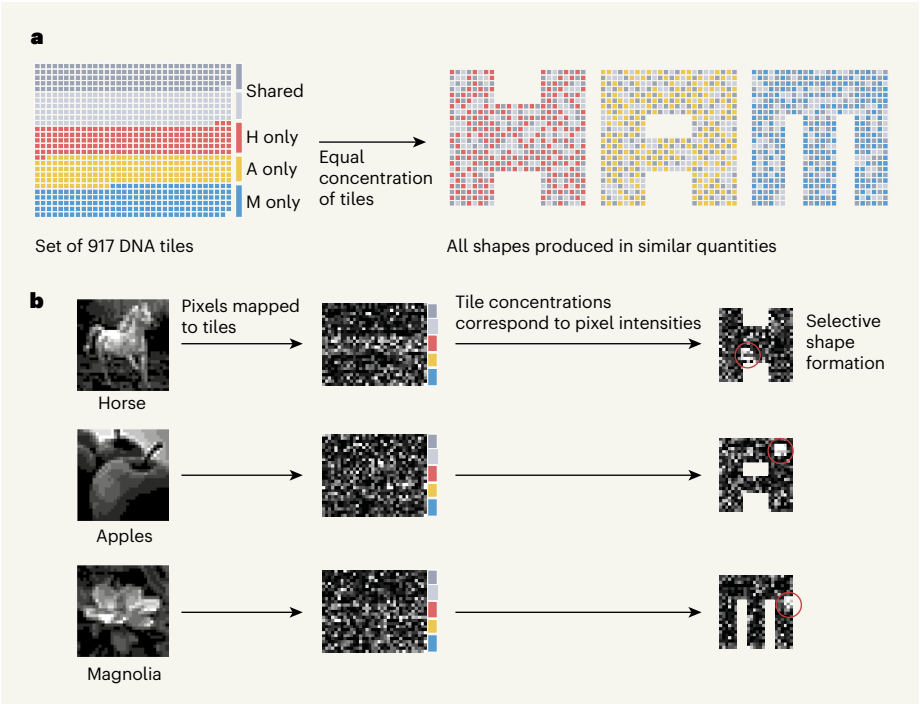

Figure 1 | Pattern recognition by self-assembling DNA tiles. **a**, Evans *et al.*[4] report a set of 917 DNA tiles that can self-assemble in solution to form three shapes: H, A and M. Some tiles (red, yellow and blue) are used in only one of the shapes, whereas others (grey) are used in two or three shapes. When the concentrations of each tile in solution are equal, all three shapes form in similar quantities. **b**, To demonstrate the pattern-recognition ability of the tile system, each pixel in 18 greyscale images (only three are shown) was assigned to one of the tiles, using the same assignment for all images. The intensity of the pixels determined the concentration of the associated tile to be used in solution: bright pixels correspond to high concentrations. The system was designed so that the tiles self-assemble predominantly into one shape that depends on the specific combination of tile concentrations defined by the images. High-concentration tiles are situated close together in the dominant shape (red circles), but not in the other shapes (not shown). This promotes the initial assembly of the high-concentration tiles, accelerating the formation of the dominant shape. Adapted from Fig. 5 in ref. 4; Magnolia: David C. Richardson.

that shape. When one of these cores forms in solution, it is called a nucleation seed, and the assembly of the rest of the shape is accelerated at the expense of the other two shapes – a 'winner take all' outcome. This means that the shape that self-assembles from the full set of tiles can be controlled by choosing a particular combination of tile concentrations, called a concentration pattern, that promotes the formation of the nucleation seed for that shape.

Evans *et al*. used an algorithm to estimate the self-assembly rates of shapes from different concentration patterns, and thereby identified 37 nucleation seeds. The corresponding concentration patterns were then tested experimentally, using a 150-hour annealing process in which the tiles were slowly cooled in solution to allow them to self-assemble. The authors used fluorescent labels to monitor the assembly of each shape, and atomic force microscopy to image the shapes. Of the 37 proposed nucleation seeds, roughly half resulted in selective self-assembly of the desired shapes; the others did not, for unknown reasons.

To demonstrate the information-processing capabilities of their system, the authors used it to classify 18 greyscale images of 30 × 30 pixels on the basis of the shades (greyscale values) of the pixels in the image (Fig. 1b). The idea was to represent the greyscale value of each pixel in each image by the concentration of one of the tiles, so that the resulting concentration pattern of tiles promotes the assembly of a designated shape (H, A or M). The assignment of pixels to tiles was done computationally to maximize the self-assembly of the designated shape, while minimizing self-assembly of competing shapes. Crucially, this assignment was simultaneously optimized for all images, rather than independently for each one.

When the authors tested the concentration patterns for the 18 images experimentally, they observed that the desired shape did indeed assemble more often than any other, with greater than 80% selectivity for 13 of the images. In other words, the tile system recognized the different concentration patterns, and therefore the corresponding images, by assembling into the designated shapes. Crucially, the system also coped with 12 degraded versions of the images. For example, when the greyscale values of some of the pixels of a horse image were altered at random, thereby corrupting the corresponding concentration pattern, the system still reliably formed an H shape in preference to an A or an M, correctly classifying the image.

One of the main limitations of this work is the trade-off between the speed, accuracy and complexity of pattern recognition. In particular, the timescales of the experiments were chosen conservatively to minimize the formation of incorrect structures. The winner-take-all outcomes indicate that these timescales could be shortened substantially; the use of smaller assemblies consisting of fewer tiles could also speed things up. Because such DNA systems will probably find biological applications, rather than becoming replacements for silicon-based computations, speed considerations might be less important than having the ability to embed computations directly in biophysical processes at the nanoscale.

Overall, the latest findings demonstrate how computations needed for complex pattern recognition can be encoded at the molecular level in the biophysical process of self-assembly. The study also illustrates how previous theoretical[7] and experimental work[6] on DNA-tile assembly can support the design of sophisticated new experiments. Furthermore, it demonstrates how the programmability of DNA and the well-understood kinetics and thermodynamics of DNA base pairing can enable the design of a self-assembling system with more than 900 distinct components to carry out complicated computations.

From a computational perspective, a promising direction for future work is to further explore the connections between pattern recognition in self-assembling systems and other forms of neural computation. Evans *et al.* identify parallels between their tile system and neural network models known as Hopfield associative memories[8], and with the networks of place cells in the brain that store spatial memories, building on previous work[8,9]. Further exploration of the opportunities and limitations of embedding neural computation in biophysical processes would be valuable.

From an experimental perspective, as scientists' ability to design protein-based systems and predict their biophysical interactions continues to improve, the approaches outlined by Evans *et al.* could be used to design self-assembling protein structures that process information. For example, different protein complexes with distinct functions could self-assemble depending on the concentrations of their building blocks. More generally, this work also provides a conceptual and experimental framework for the future design of compact, robust and scalable computations embedded in biophysical processes.

Andrew Phillips is at Sanofi, Babraham Research Campus, Cambridge CB22 3AT, UK.
e-mail: andrew.phillips@sanofi.com

1. Bray, D. *Nature* **376**, 307–312 (1995).
2. Dayan, P. & Abbott, L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, 2005).
3. Cohen, I. R. & Efroni, S. *Front. Immunol.* **10**, 10 (2019).
4. Evans, C. G., O'Brien, J., Winfree, E. & Murugan, A. *Nature* **625**, 500–507 (2024).
5. Seeman, N. C & Sleiman, H. F. *Nature Rev. Mater.* **3**, 17068 (2018).
6. Woods, D. *et al. Nature* **567**, 366–372 (2019).
7. Winfree, E. & Bekbolatov, R. in *DNA Computing* (Lecture Notes in Computer Science) Vol. 2943 (eds Chen, J. & Reif, J.) 126–144 (Springer, 2004).
8. Murugan, A., Zeravcic, C, Brenner, M. P. & Leibler, S. *Proc. Natl Acad. Sci. USA* **112**, 54–59 (2015).
9. Zhong, W., Schwab, D. J. & Murugan, A. *J. Stat. Phys.* **167**, 806–826 (2017).

The author declares no competing interests.

Biomedical materials

Composite gels designed to stick to biological tissue

Sophia J. Bailey & Eric A. Appel

Materials that adhere tightly to human tissues can promote healing and boost the sensitivity of biomedical diagnostic devices. An 'evolving' gel has been made that synergizes two strategies for forming interfaces with tissue.

The interfaces that form between living tissues and biomedical materials, often referred to as biointerfaces, greatly influence the ability to detect and treat disease. Unfortunately, the development of biomedical devices has historically involved a trade-off between using materials that can be fabricated easily into the devices and using those that adhere to tissues at the cellular level. Reporting in *Nature Chemical Engineering*, Shi *et al.*[1] present a clever, yet simple, strategy to make materials that combine easy handling with robust interface formation. The authors show that hydrogels – water-rich networks of polymers – embedded with tiny starch granules form dynamic biointerfaces that 'evolve' over time, and have many potential uses across biomedicine, from tissue regeneration to sensing.

Materials that are intended to help repair damaged tissues must provide a tight biointerface to promote cell–material interactions, and have a porous or degradable matrix to allow for cellular growth[2]. Typically, macroscopic patches are used to promote such