DNA Computing

by Miriam Balaban
Boston University, Center for Philosophy
and History of Science
(Based on articles in references)

Whereas current technology rests on a highly linear principle of logic, the use of DNA means that an enormous number of calculations can take place simultaneously.

DNA's ability to store and process information

(From EMBO Report, 2003. 4: 7–10.)

Research into DNA has given biologists a great understanding of life, and has also allowed them to create innumerable useful tools that have wide-ranging applications. However, it was not until the early 1990s that researchers started exploring the possibility of utilizing DNA's ability to store and process information outside the realms of biology. In 1994, a study showed that DNA could be used to solve mathematical problems, which attracted considerable interest from researchers hoping that DNA would one day replace silicon as the basis for a new wave of computers. But the initial excitement has since dampened down as scientists have realized that there are numerous problems inherent to DNA computing and that they would have to live with their silicon-based computers for quite a while yet. The field consequently changed its focus, and in essence, research into DNA computing is now chiefly concerned with investigating processes in cells that can be viewed as logical computations and then looking to use these computations to our advantage. A mix of 10^{18} strands of DNA could operate at 10,000 times the speed of today's advanced supercomputers

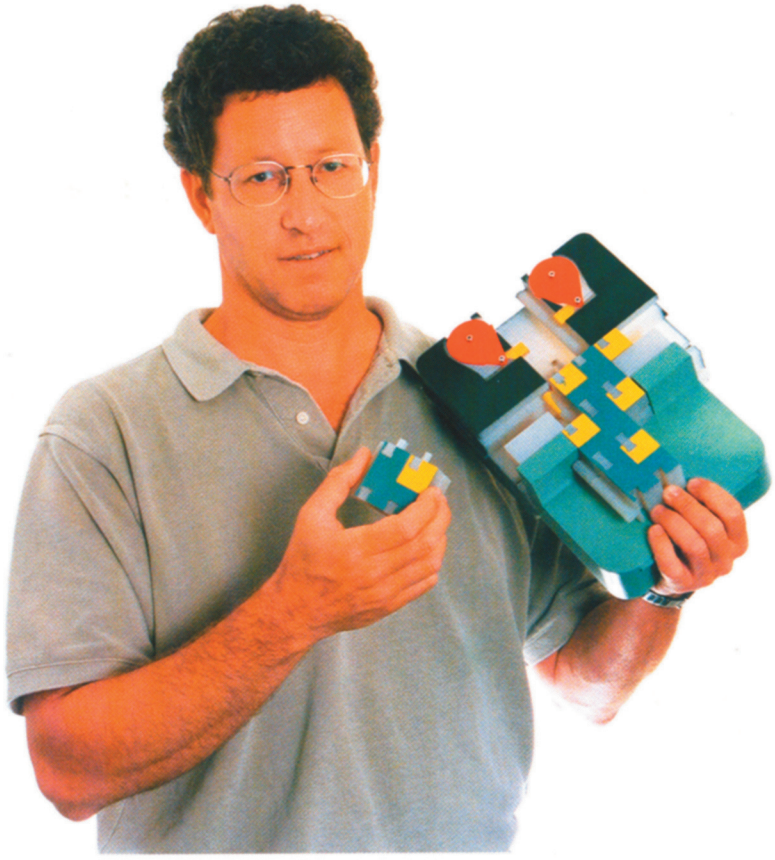
It was Leonard Adleman, Professor of Computer Science and Molecular Biology at the University of Southern California, USA, who in 1994 pioneered the field when he built the first DNA based computer. Intrigued by the molecule's immense capacity to store information in a very small space, he set out to solve a classic puzzle in mathematics — the so-called Hamilton Path problem, better known as the Traveling Salesman problem. This seemingly simple puzzle — a salesman must visit a number of cities that are interconnected by a limited series of roads without passing through any city more than once – is actually quite a killer, and even the most advanced supercomputers would take years to calculate the optimal route for 50 cities. Adleman solved the problem for seven cities within a second, using DNA molecules in a standard reaction tube. He represented each of the seven cities as separate, single-stranded DNA molecules, 20 nucleotides long, and all possible paths between cities as DNA molecules composed of the last ten nucleotides of the departure city and the first ten nucleotides of the arrival city. Mixing the DNA strands with DNA ligase and adenosine triphosphate (ATP) resulted in the generation of all possible random paths through the cities. However, the majority of these paths were not applicable to the situation–they were either too long or too short, or they did not start or finish in the right city. Adleman then filtered out all the paths that neither started nor ended with the correct molecule and those that did not have the correct length and composition. Any remaining DNA molecules represented a solution to the problem. The power contained in these tiny molecules caused a flurry of excitement in the computing world

The computation in Adleman's experiment chugged along at 10^{14} operations per second, a rate of 100 Teraflops or 100 trillion floating point operations per second; the world's fastest supercomputer, Earth Simulator, owned by the NEC Corporation in Japan, runs at just 35.8 Teraflops. Clearly, computing with DNA has massive advantages over silicon-based machines. Whereas current technology rests on a highly linear principle of logic, and one computation must be completed before the next can begin, the use of DNA means that an enormous number of calculations can take place simultaneously. This parallel power is many times faster than that of traditional machines — a mix of 1018 strands of DNA could operate at 10,000 times the speed of today's advanced supercomputers. The other major advantage is the potential for information storage. Whereas traditional storage media, such as

videotapes, require 10^{12} cubic nanometers of space to store a single bit of information, DNA molecules require just one cubic nanometer per bit. Not surprisingly, the power contained in these tiny molecules caused a flurry of excitement in the computing world, and many hoped that DNA computing could overtake silicon-based technology.

Although the potential of DNA computing seemed enormous, intervening research has shown that it is constrained by major limitations. The emphasis has now shifted away from the original objective. There is still great potential in DNA computing but the rich potential of DNA computing lies in *in vivo* computing" — using the technology on a smaller scale, inside cells.

A demonstration of this aim was achieved in 2001 by Ehud Shapiro's group at the Weizmann Institute of Science in Israel who built a programmable and autonomous computing machine made of biomolecules.



Prof. Shapiro holding his design of a mechanical Turing machine which served as an inspiration for the molecular automaton.

MOLECULAR AUTOMATA

A trillion computers in a drop of water: building a nanoscale computing machine using biological molecules

Benenson, Y., Paz-Elizur, T., Adar, R., Keinan, E., Livneh, Z. Shapiro, E. (2001) Nature, *414: 430-434 Weizmann Media Relations Department*

Prof. Ehud Shapiro and his colleagues are using biological molecules to create a tiny computer a programmable two-state, two-symbol finite automaton in a test tube. This biological nanocomputer is so small that a trillion (1,000,000,000,000) such computers co-exist and compute in parallel, in a drop the size of 1/10 of a milliliter of watery solution held at room temperature. Collectively, the computers perform a billion operations per second with greater than 99.8% accuracy per operation while requiring less than a billionth of a Watt of power. This study may lead to future computers that can operate within the human body, interacting with its biochemical environment to yield far-reaching biological and pharmaceutical applications.

The computer's input, output, and 'software' are made up of DNA molecules. For 'hardware,' the computer uses two naturally occurring enzymes that manipulate DNA. When mixed together in solution, the software and hardware molecules operate in harmony on the input molecule to create the output molecule, forming a simple mathematical computing machine, known as a finite automaton. This nanocomputer can be programmed to perform several simple tasks by choosing different software molecules to be mixed in solution. For instance, it can detect whether, in an input molecule encoding a list made of 0's and 1's, all the 0's precede all the 1's.

'The living cell contains incredible molecular machines that manipulate information-encoding molecules such as DNA and RNA in ways that are fundamentally very similar to computation,' says Prof. Shapiro of

A spoonful of "computer soup" can contain 15,000 trillion computers, together performing 330 trillion operations per second

the Institute's Computer Science and Applied Mathematics Department and the Biological Chemistry Department. 'Since we don't know how to effectively modify these machines or create new ones just yet, the trick is to find naturally existing machines that, when combined, can be steered to actually compute.'

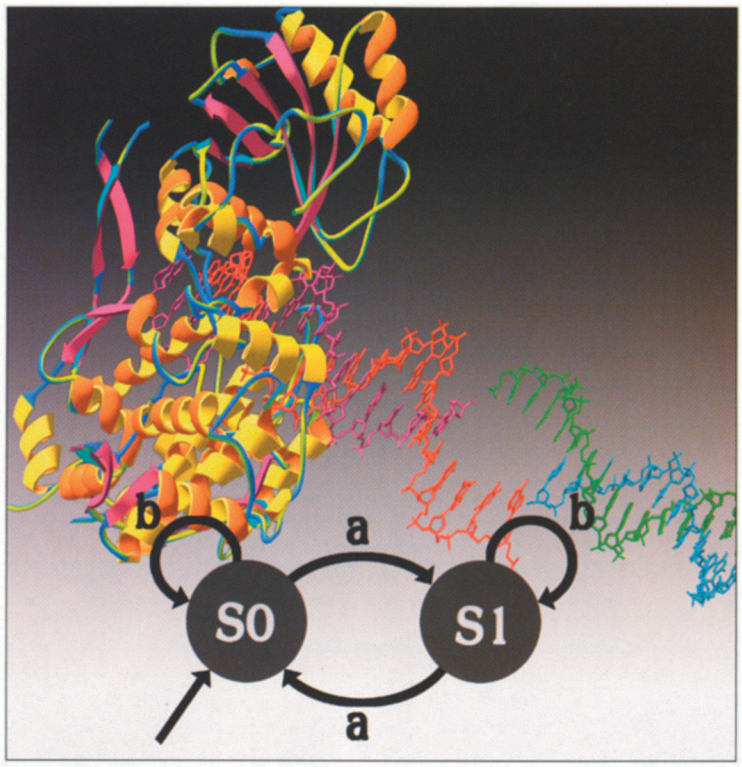
Shapiro challenged his Ph.D. student, Yaakov Benenson, to do just that: to find a molecular realization of one the simplest mathematical computing machines a finite automaton that detects whether a list of 0's and 1's has an even number of 1's. Benenson came up with a solution using DNA molecules and two naturally occurring DNA-manipulating enzymes: *Fok*I and Ligase. Operating much like a biological editing kit, *Fok*I functions as a chemical scissors, cleaving DNA in a specific pattern, whereas the Ligase enzyme seals DNA molecules together.

As the lab work progressed, Shapiro and his team realized that the automaton they built could be programmed to perform different tasks by selecting different subsets of the molecules realizing the eight possible rules of operation controlling the performance of a two-state, two-symbol finite automaton. The software molecules, together with two 'output display' molecules used to visualize the final result of the computation, can be used to create a total of 765 software programs. Several of these programs were tested in the lab, including the 'even 1's checker' and the '0's before 1's' test mentioned above, as well as programs that check whether a list of 0's and 1's has at least (or at most) one 0, and whether it both starts with a 0 and ends with a 1.

The nanocomputer created by Shapiro's team uses the four DNA bases known as A, G, C and T, to encode the input data as well as the program rules underlying the computer 'software.' Both input and software molecules are designed to have one DNA strand longer than the other, resulting in a single-strand overhang called a 'sticky end.' Two molecules with complementary sticky ends can temporarily stick to each other (a process known as hybridization), allowing DNA Ligase to permanently seal them into one molecule. The sticky end of the input molecule encodes the current symbol and the current state of the computation, whereas the sticky end of each 'software' molecule is designed to detect a particular state-symbol combination. A two-state, two-symbol automaton has four such combinations. For each combination the nanocomputer has two possible next moves, to remain in the same state or to change to the other state, allowing eight software molecules to cover all possibilities.

In each processing step the input molecule hybridizes with a software molecule that has a complementary sticky end, allowing Ligase to seal them together using two ATP molecules as energy. Then comes *Fok*I, detecting a special site in the software molecule known as the recognition site. It cleaves the input molecule in a location determined by the software molecule, thus exposing a sticky end that encodes the next input symbol and the next state of the computation. Once the last input symbol is processed, a sticky end encoding the final state of the computation is exposed and detected, again by hybridization and ligation, by one of two 'output display' molecules. The resulting molecule, which reports the output of the computation, is made visible to the human eye in a process known as gel electrophoresis.

The nanocomputer created is too simple to have immediate applications. However it may pave the way to future computers that can operate within the human body with unique biological and pharmaceutical applications. 'For instance, such a future computer could sense an abnormal biochemical change in the body and decide how to correct it by synthesizing and releasing the necessary drug,' says Prof. Zvi Livneh, a DNA expert



Bottom: A finite state automaton that analyzes its input, a list of a s and bs, according to the program rules depicted as labeled arrows, ending the computation in S0 if the input has an even number of a s and in S1 otherwise.

Top: Molecular realization of this automaton. An input DNA molecule (green/blue) provides both data and fuel for the computation. Software DNA molecules (red/purple) encode pogram rules, and the restriction enzyme *FokI* (colored ribbons) functions as the automaton's hardware. (Image source: Proceedings of the National Academy of Sciences).

from the Institute's Department of Biological Chemistry who collaborated on this project.

Other contributors to this research include Dr. Tamar Paz-Elizur and Dr. Rivka Adar from the Weizmann Institute's Department of Biological Chemistry, and Prof. Ehud Keinan from the Department of Chemistry at the Technion – Israel Institute of Technology, Haifa and the Department of Molecular Biology of The Scripps Research Institute, La Jolla, CA

DNA fuels molecular computer: computer awarded Guinness World Record

Benenson, Y., Adar, R., Paz-Elizur, T., Livneh, Z. Shapiro, E. (2003) Proc. Natl. Acad. Sci. USA *(2003) 100: 2191-2196.*

Ehud Shapiro and his team have found another new use for DNA: fuel for molecular computation systems.

Whether plugged in or battery-powered, computers need energy. Around a year ago, Ehud Shapiro made international headlines for devising a programmable molecular computing machine composed of enzymes and DNA molecules. Now his team has made the device uniquely frugal: the single DNA molecule that provides the computer with the input data also provides all the necessary fuel.

The source of fuel of the earlier device was a molecule called ATP, the standard energy currency of all life forms. The redesigned device processes its DNA input molecule using only spontaneous, energy-releasing operations. It breaks two bonds in the DNA input molecule, releasing the energy stored in these bonds as heat. This process generates sufficient energy to carry out computations to completion without any external source of energy.

A spoonful (5 milliliters) of "computer soup" can contain 15,000 trillion such computers, together performing 330 trillion operations per second with 99.9% accuracy per step. These computers need very little energy (all supplied, as mentioned, by the input molecule) and together release less than 25 millionths of a watt as heat.

The device was recently awarded the Guinness World Record for "smallest biological computing device."

The study was carried out by Yaakov Benenson, Dr. Rivka Adar, Dr. Tamar Paz-Elizur, Prof. Zvi Livneh and Prof. Ehud Shapiro of the Institute's Biological Chemistry Department and the Computer Science and Applied Mathematics Department. ■

References

Adelman L.M., (1994) *Science* **266**: 1021–1022.

Benenson, Y., Paz-Elizur, T., Adar, R., Keinan, E., Livneh, Z. Shapiro, E. (2001) *Nature*, **414**: 430-434.

Benenson, Y., Adar, R., Paz-Elizur, T., Livneh, Z. Shapiro, E. (2003) *Proc. Natl. Acad. Sci. USA* (2003) **100**: 2191-2196

Parker J, (2003) *EMBO Report*, **4**: 7–10

The Weizmann Institute of Science Media Information.

A video presentation in beta format illustrating the operation of the molecular computer developed by Prof. Shapiro's team is available upon request.

Ehud.Shapiro@weizmann.ac.il