Solution of a 20-Variable 3-SAT Problem on a DNA Computer

Ravinderjit S. Braich,[1] Nickolas Chelyapov,[1] Cliff Johnson,[1] Paul W. K. Rothemund,[2] Leonard Adleman[1]*

A 20-variable instance of the NP-complete three-satisfiability (3-SAT) problem was solved on a simple DNA computer. The unique answer was found after an exhaustive search of more than 1 million (2^{20}) possibilities. This computational problem may be the largest yet solved by nonelectronic means. Problems of this size appear to be beyond the normal range of unaided human computation.

The vast parallelism, exceptional energy efficiency, and extraordinary information density inherent in molecular computation have raised the possibility that molecular computers might some day prove capable of attacking problems that have resisted conventional methods (*1–6*).

Numerous architectures for molecular computing have been proposed (*1, 2, 7–17*), and of these, several have been explored experimentally and found to be feasible (*1, 6, 8, 13, 14, 16–19*).

The 3-SAT problem is an NP-complete computational problem (*20*) for which the fastest known sequential algorithms require exponential time. The problem became the benchmark for testing the performance of DNA computers, after Lipton (*2*) demonstrated that it was well suited to take advantage of the parallelism afforded by molecular computation. A group led by Smith (*14*) used surface-based chemistry to solve a four-variable (16 possible truth assignments) instance of the problem. Yoshida and Suyama (*18*) also solved a four-variable instance using a DNA program implementing a breadth first search. Sakamoto *et al.* (*13*) solved a six-variable (64 possible truth assignments) problem using hairpin DNA. A group led by Landweber (*16*) used RNA to solve an instance of a nine-variable (512 possible truth assignments) satisfiability problem related to the "Knights Problem" in chess. Here, a 20-variable (1,048,576 possible truth assignments) instance of the 3-SAT problem is solved.

In the present study, the architecture employed is related to the Sticker Model described by Roweis *et al.* (*9*). The Sticker Model uses two basic operations for computation: separation based on subsequence and application of stickers. Only separations are used in the current study. Separations are carried out using oligonucleotide probes immobilized in polyacrylamide gel–filled glass modules. Information-carrying DNA strands are moved through the modules by electrophoresis. Strands with subsequences complementary to those of the immobilized probes hybridize and are retained in the module; strands without complementary subsequences pass through the module relatively unhindered (*21*). Captured strands are released from the probes by running electrophoresis at a temperature higher than the melting temperature of the probe/complement duplex. Released strands may be transported via electrophoresis to new modules for further separations.

Using electrophoresis to transport DNA strands between gel-filled glass modules results in a computer that is "dry" and potentially automatable. Because covalent bonds are neither made nor broken during separations, DNA strands and glass modules are potentially reusable for multiple computations.

The Boolean formula. The input to the computation was a 20-variable 24-clause 3-conjunctive normal form (3-CNF) formula, Φ (Fig. 1A). To make the computation as challenging as possible, Φ was designed to have a unique satisfying truth assignment (Fig. 1B). This design has endowed Φ with an iterative syntactical structure that can be seen on close inspection. It is important to note, however, that the DNA computation undertaken here made no use of this structure, and instead, exhaustively searched all 2^{20} (1,048,576) possible truth assignments in the process of finding the unique satisfying assignment. Hence, it is reasonable to assume that any 20-variable 24-clause 3-CNF formula would have been just as readily solved.

The library. To represent all possible truth assignments, a Lipton encoding (*2*) was used. For each of the 20 variables x_k ($k = 1 \ldots 20$), two distinct 15 base "value sequences" were designed: one representing true (T), X^T_k, and one representing false (F), X^F_k (*22*). For $k = 1 \ldots 20$, $Z = T$ or F, $\bar{X}^Z_k$ will denote the Watson-Crick complement of X^Z_k. Each of the 2^{20} truth assignments was represented by a "library sequence" of 300 bases consisting of the ordered concatenation of one value sequence for each variable. Single-stranded DNA molecules with library sequences were termed "library strands." A collection of all library strands duplexed with complements was termed a "full library."

For each of the 40 sequences $\bar{X}^Z_k$, $k = 1, \ldots 20$, $Z = T$ or F, 5′-end Acrydite-modified (Mosaic Technologies, Boston, MA) oligonucleotides were obtained [Operon Technology, Alameda, CA, or Integrated DNA Technolo-

A

$\Phi = (\sim x_3 \text{ or } \sim x_{16} \text{ or } x_{18})$ *and* $(x_5 \text{ or } x_{12} \text{ or } \sim x_9)$ *and* $(\sim x_{13} \text{ or } \sim x_2 \text{ or } x_{20})$ *and* $(x_{12} \text{ or } \sim x_9 \text{ or } \sim x_5)$ *and* $(x_{19} \text{ or } \sim x_4 \text{ or } x_6)$ *and* $(x_9 \text{ or } x_{12} \text{ or } \sim x_5)$ *and* $(\sim x_1 \text{ or } x_4 \text{ or } \sim x_{11})$ *and* $(x_{13} \text{ or } \sim x_2 \text{ or } \sim x_{19})$ *and* $(x_5 \text{ or } x_{17} \text{ or } x_9)$ *and* $(x_{15} \text{ or } x_9 \text{ or } \sim x_{17})$ *and* $(\sim x_5 \text{ or } \sim x_9 \text{ or } \sim x_{12})$ *and* $(x_6 \text{ or } x_{11} \text{ or } x_4)$ *and* $(\sim x_{15} \text{ or } \sim x_{17} \text{ or } x_7)$ *and* $(\sim x_6 \text{ or } x_{19} \text{ or } x_{13})$ *and* $(\sim x_{12} \text{ or } \sim x_9 \text{ or } x_5)$ *and* $(x_{12} \text{ or } x_1 \text{ or } x_{14})$ *and* $(x_{20} \text{ or } x_3 \text{ or } x_2)$ *and* $(x_{10} \text{ or } \sim x_7 \text{ or } \sim x_8)$ *and* $(\sim x_5 \text{ or } x_9 \text{ or } \sim x_{12})$ *and* $(x_{18} \text{ or } \sim x_{20} \text{ or } x_3)$ *and* $(\sim x_{10} \text{ or } \sim x_{18} \text{ or } \sim x_{16})$ *and* $(x_1 \text{ or } \sim x_{11} \text{ or } \sim x_{14})$ *and* $(x_8 \text{ or } \sim x_7 \text{ or } \sim x_{15})$ *and* $(\sim x_8 \text{ or } x_{16} \text{ or } \sim x_{10})$

B

x_1=F, x_2=T, x_3=F, x_4=F, x_5=F, x_6=F, x_7=T, x_8=T, x_9=F, x_{10}=T, x_{11}=T, x_{12}=T, x_{13}=F, x_{14}=F, x_{15}=T, x_{16}=T, x_{17}=T, x_{18}=F, x_{19}=F, x_{20}=F

Fig. 1. The computational problem. **(A)** 20-variable 3-CNF Boolean formula Φ. The symbol "~" indicates "not." **(B)** The unique truth assignment satisfying Φ.

[1]University of Southern California, Laboratory for Molecular Science, Los Angeles, CA 90089–1340, USA. [2]California Institute of Technology, Pasadena, CA 91125, USA.

*To whom correspondence should be addressed. E-mail: Adleman@USC.edu. Web site: www.usc.edu/dept/molecular-science/index.html

gies (IDT), Skokie, IL] and used as probes during separation operations.

To reduce errors in computation, sequences were designed to discourage intra- and interlibrary strand hybridization and unintended probe-library strand hybridization. To achieve these goals, sequences were computer generated to satisfy previously reported constraints (*19*). In particular, G's do not occur in value sequences.

Synthesis of long molecules on automated DNA synthesizers can be inefficient. To avoid such inefficiencies, creation of the full library began with the synthesis of two "half libraries," one for variables x_1 through x_{10} (the left half-library) and one for variables x_{11} through x_{20} (the right half-library). Half libraries were synthesized on a dual column ABI 392 DNA/RNA Synthesizer (Applied Biosystems, Foster City, CA) at a 0.2-μmol scale using polystyrene-based solid support. A mix-and-split combinatorial synthesis technique was used (*16*). Briefly, for the left half-library, oligonucleotides with sequences X^T_{10} and X^F_{10} were synthesized in separate columns. The columns were removed from the synthesizer and opened. The solid-support beads were mixed together and divided into halves, which were then put into separate columns. The columns were closed and synthesis restarted with sequences X^T_9 and X^F_9 in separate columns. This process was repeated until all 10 variables had been processed. A similar process was used for the right half-library.

To assess the degeneracy of the half-libraries and the efficacy of capturing half-libraries with Acrydite-modified probes, a gel capture experiment was performed [Web fig. 1 (*23*)]. For each of the 40 sequences $\bar{X}^Z_k$, $k = 1 \ldots 20$, $Z = T$ or F, a "capture layer" was created by adding the corresponding Acrydite-modified probe to polyacrylamide gel. An aliquot of the appropriate 5′-[^{32}P]labeled half-library was run through the capture layer via electrophoresis. As expected, for each of the 40 probes, approximately half of the strands of the half-library were captured whereas approximately half passed through. This suggested that probes stayed in the gels and captured half-library strands. This also suggested that half-library strands had subsequences complementary to those of the probes and that, for each variable, the number of half-library strands representing true and the number representing false were approximately equal.

To further test the half-libraries, polymerase chain reaction (PCR) amplifications were run with primer sets: $<X^T_1, \bar{X}^T_{10}>$, $<X^T_1, \bar{X}^F_{10}>$, $<X^F_1, \bar{X}^T_{10}>$, $<X^F_1, \bar{X}^F_{10}>$, under standard conditions (*24*) for 35 cycles using 400 fmol of left half-library as template. Similarly, PCR amplifications were run with primer sets: $<X^T_{11}, \bar{X}^T_{20}>$, $<X^T_{11}, \bar{X}^F_{20}>$, $<X^F_{11}, \bar{X}^T_{20}>$, $<X^F_{11}, \bar{X}^F_{20}>$, under standard conditions for 35 cycles using 400 fmol of right half-library as template. Gel analysis showed that products of expected lengths were obtained in all cases (*25*). This confirmed that subsequences X^T_1, X^F_1, X^T_{10}, X^F_{10} were present at the expected positions in the left half-library, and that subsequences X^T_{11}, X^F_{11}, X^T_{20}, X^F_{20} were present at the expected positions in the right half-library.

The 300-oligomer (300-mer) full library was created from the two half-libraries using a polymerase extension method similar to that described in (*26*). Subsequently, two stages of PCR amplification were performed to produce the quantity of full library required for the computation. Gel analysis of the final product showed only the presence of a band corresponding to 300 base pairs (bp) (Fig. 2). Spectrophotometric analysis of the full library showed a total of approximately 750 pmol of 300-bp DNA. Note that the full library consisted of library strands duplexed with Acrydite-modified complements. Because all library strands had the same length (300 bases), they were expected to run at similar rates during electrophoresis.

To test the full library, PCR amplifications were run with primer sets: $<X^T_1, \bar{X}^T_k>$, $<X^T_1, \bar{X}^F_k>$, $<X^F_1, \bar{X}^T_k>$, $<X^F_1, \bar{X}^F_k>$, $<X^T_1, X^F_1, \bar{X}^T_k, \bar{X}^F_k>$ for various k. Gel analysis of the resulting products showed bands of the expected lengths (Fig. 3).

The computer and the computational protocol. The computer consisted of an electrophoresis box with a hot chamber and a cold chamber, a glass "library module" filled with polyacrylamide gel containing covalently bound full library, and for each of the 24 clauses of Φ, a glass "clause module" filled with polyacrylamide gel containing covalently bound probes and designed to capture only library strands encoding truth assignments satisfying that clause (Fig. 4). The computational protocol was as follows:

Step 1: Insert the library module into the hot chamber of the electrophoresis box and the first clause module into the cold chamber of the box. Begin electrophoresis. In theory, during Step 1, library strands melt off their Acrydite-modified complements in the library module and migrate to the first clause module. Library strands encoding truth assignments satisfying the first clause are captured in the capture layer, while library strands encoding nonsatisfying assignments

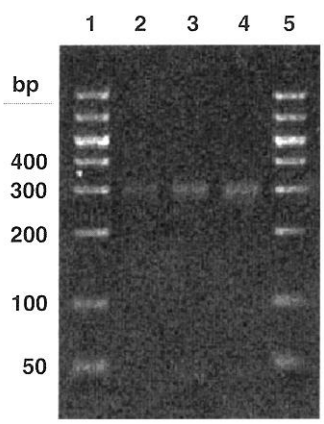

Fig. 2. Formation of the full-library. Four 30-mer oligonucleotide "splints" were obtained (IDT) with sequences: $\bar{X}^T_{11}\bar{X}^T_{10}$, $\bar{X}^T_{11}\bar{X}^F_{10}$, $\bar{X}^F_{11}\bar{X}^T_{10}$, $\bar{X}^F_{11}\bar{X}^F_{10}$. Ten pmol of each half-library was mixed with 2 pmol of each of the four splints, in a final volume of 25 μl in 1× T4 DNA ligase buffer (Promega, Madison, WI) (note: no ligase was added) and incubated at room temperature for 2 hours. One-half microliter of the mixture was PCR-amplified with primer set: $<X^T_1, X^F_1$, Acrydite-modified $\bar{X}^T_{20}$, Acrydite-modified $\bar{X}^F_{20}>$, under standard conditions for 35 cycles. Then, 1-μl aliquots were PCR-amplified from the product above, using the above primer set, under standard conditions for 35 cycles. The products were run on a 1% agarose gel, the 300-bp band extracted with Ultrafree-DA DNA extraction kit (Millipore, Bedford, MA), then pooled to create a stock solution with a final volume of ~500 μl. Another round of PCR amplification was performed using the above primer set under standard conditions for 15 cycles with 2.5-μl aliquots of stock solution as template. The resulting DNA was ethanol-precipitated and rehydrated in 75 μl of water. Shown here are 1 μl (lane 2), 2 μl (lane 3), and 3 μl (lane 4) aliquots of the full library run on a 4% agarose gel. Lanes 1 and 5 are molecular weight markers.

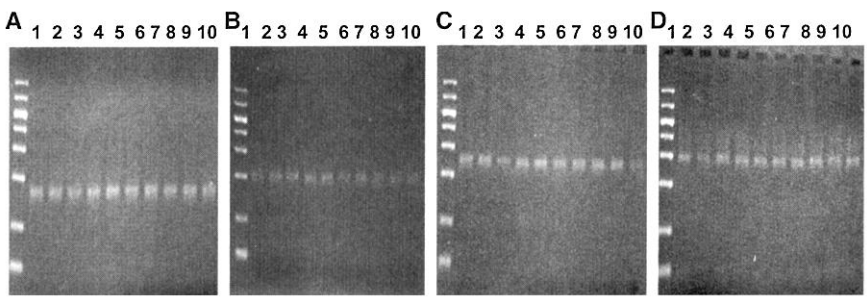

Fig. 3. Analysis of the full library. Purified full library was PCR-amplified under standard conditions for 15 cycles. PCR products were analyzed on 4% agarose gels. Lanes 1 and 2 correspond to primer set $<X^T_1, X^F_1, \bar{X}^T_k, \bar{X}^F_k>$, lanes 3 and 4 correspond to primer pair $<X^T_1, \bar{X}^T_k>$, lanes 5 and 6 correspond to primer pair $<X^T_1, \bar{X}^F_k>$, lanes 7 and 8 correspond to primer pair $<X^F_1, \bar{X}^T_k>$, lanes 9 and 10 correspond to primer pair $<X^F_1, \bar{X}^F_k>$, where (**A**) $k = 11$; (**B**) $k = 14$; (**C**) $k = 17$; (**D**) $k = 20$. Molecular weight markers (as in Fig. 2) are on the leftmost lane of each gel.

run through the capture layer and continue into the buffer reservoir. In particular, those library strands with sequences X^F_3 or X^F_{16} or X^T_{18} are retained in the capture layer, whereas those with sequences X^T_3 and X^T_{16} and X^F_{18} run through.

Step 2: Remove both modules from the box. Discard the module from the hot chamber. Wash the box and add new buffer. Insert the module from the cold chamber into the hot chamber and the module for the next clause into the cold chamber. Begin electrophoresis. In theory, during Step 2, library strands melt off their Acrydite-modified probes in the clause module located in the hot chamber and migrate to the clause module in the cold chamber. Library strands encoding truth assignments satisfying the clause associated with the module in the cold chamber will be captured, while library strands encoding non-satisfying assignments will run through the capture layer and continue into the buffer reservoir.

Step 3: Repeat Step 2 for each of the remaining 22 clauses. In theory, at the end of Step 3, the final (24th) clause module will contain only those library strands which have been captured in all 24 clause modules and hence encode truth assignments satisfying each clause of Φ and therefore Φ itself.

Step 4: Extract the answer strands from the final clause module, PCR-amplify, and "read" the answer.

Determination of the answer. Gel was extruded from the final (24th) clause module and soaked in 1 ml of water at 65°C overnight to extract the library strands it contained. The library strands were lyophilized, reconstituted in 200 μl of water, desalted, and recovered in 45 μl of water. This became the answer stock.

For assigning truth values to variables x_1 and x_{20}, 1-μl aliquots of 10-, 20-, 30-, 40-, 50-, 60-, and 100-fold dilutions of the answer stock were PCR-amplified with primer sets: $\langle X^T_1, \bar{X}^T_{20}\rangle$, $\langle X^T_1, \bar{X}^F_{20}\rangle$, $\langle X^F_1, \bar{X}^T_{20}\rangle$, $\langle X^F_1, \bar{X}^F_{20}\rangle$. Gel analysis of the PCR products for 10-, 20-, 30-, 40-, 50-fold dilutions showed no bands except for primer set: $\langle X^F_1, \bar{X}^F_{20}\rangle$. These primer sets gave only a band corresponding to 300 bp. Based on this, x_1 was assigned to be *F* and x_{20} was assigned to be *F*. Analysis of the PCR products for the 60- and 100-fold dilutions showed no bands for any primer set (*25*).

For assigning truth values to the variables $x_2, x_3 \ldots x_{19}$, and as a redundant test for the truth value of x_{20}, a 1 μl aliquot of the 50-fold dilution of the answer stock was PCR-amplified with primer sets: $\langle X^T_1, \bar{X}^T_k\rangle$, $\langle X^T_1, \bar{X}^F_k\rangle$, $\langle X^F_1, \bar{X}^T_k\rangle$, $\langle X^F_1, \bar{X}^F_k\rangle$, where k = 2, 3. . .20. Gel analysis showed (Fig. 5) that in each case only one combination of primers gave a band and this band was of the expected length (compare with Fig. 3). On this basis, truth-values were assigned to each variable. These experimentally derived truth-values corresponded to the unique satisfying truth assignment for Φ (Fig. 1B).

Capture-release efficiency. The following analysis for correct strands, those encoding the unique satisfying truth assignment, was made. Because PCR of 1 μl of a 50-fold dilution of the 45 μl answer stock revealed only the correct truth assignment, it is probable that at a minimum, 50 × 45 = 2250 correct strands were in the answer stock. Since the computation began with approximately 500 pmol (3×10^{14} molecules) of full library, and since approximately 1 in 2^{20} were correct strands, it follows that the probability of a correct strand surviving the entire computation was at least $2250 \times (2^{20})/(3 \times 10^{14}) = 7.5 \times 10^{-6}$. Because the computation had 24 capture-release steps, it follows that on average, the probability of a correct strand surviving a single capture-release step was at least $(7.5 \times 10^{-6})^{(1/24)} = 0.61$. This analysis assumes that PCR was

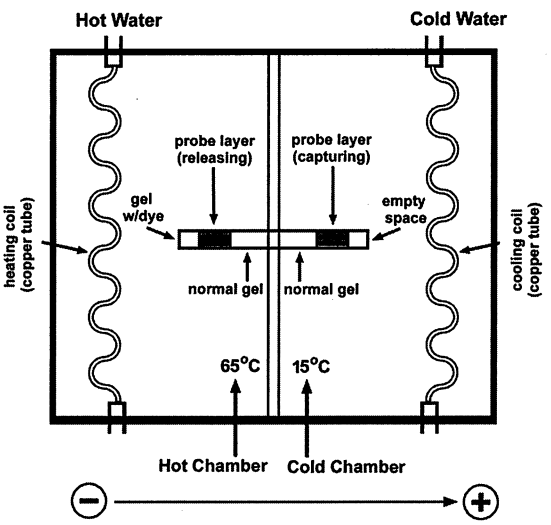

Fig. 4. The computer. An electrophoresis box 30 cm long, 15 cm wide, and 8 cm high was constructed from 0.5-cm-thick plexiglass. The box was partitioned into a hot chamber and a cold chamber of equal volumes, by a plexiglass divider. Each chamber was attached via plastic tubing to a circulating water bath. Water from the bath was transported across the chamber through copper tubing, acting as a cooling/heating coil. A platinum wire electrode was inserted into each chamber. For each of the 24 clauses of Φ, a 100-μl clause solution was prepared containing 15 μM of each of three Acrydite-modified probes, one for each literal in the clause (if x_k appeared in the clause, the probe with sequence $\bar{X}^T_k$ was added, if $\sim x_k$ appeared, the probe with sequence for $\bar{X}^F_k$ was added). For example, for the first clause, ($\sim x_3$ or $\sim x_{16}$ or x_{18}), probes with sequences $\bar{X}^F_3$, $\bar{X}^F_{16}$, and X^T_{18} were added. For each clause solution, a clause module was created in a 4.5-cm-long glass tube with outer diameter of 0.5 cm and inner diameter of 0.3 cm. A base layer of 5% polyacrylamide gel was polymerized in the bottom 3.2 cm of the tube. A probe layer of 5% polyacrylamide gel containing the clause solution was polymerized on top of the base layer. A library module was prepared as above, but instead of a clause solution, 500 pmol of the full library in 100 μl was used (*28*). The end of the module destined for the hot chamber was plugged with 5% polyacrylamide gel containing bromophenol blue and xylene cyanol dyes. The hot chamber was brought to 65°C and the cold chamber to 15°C. Electrophoresis was performed at 12 V/cm. Movement of the dyes through the gel provided a means of monitoring progress and of detecting possible leaks at the interface between the modules. Electrophoresis was stopped when the Xylene Cyanol dye had gone through the modules and into the cold buffer, approximately 4 hours.

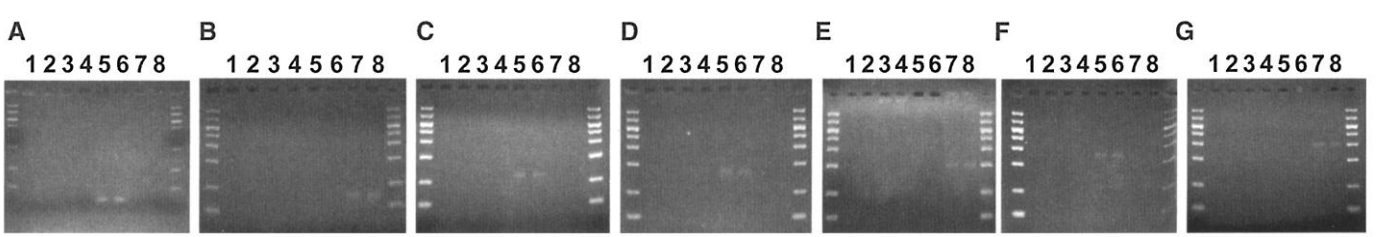

Fig. 5. Readout of the answer: 1-μl aliquots of a 50-fold dilution of the answer stock were PCR-amplified under standard conditions for 25 cycles. PCR products were analyzed on 4% agarose gels. Lanes 1 and 2 correspond to primer pair $\langle X^T_1, \bar{X}^T_K\rangle$, lanes 3 and 4 correspond to primer pair $\langle X^T_1, \bar{X}^F_K\rangle$, lanes 5 and 6 correspond to primer pair $\langle X^F_1, \bar{X}^T_K\rangle$, lanes 7 and 8 correspond to primer pair $\langle X^F_1, \bar{X}^F_K\rangle$, where (**A**) k = 2, (**B**) k = 5, (**C**) k = 8, (**D**) k = 11, (**E**) k = 14, (**F**) k = 17, (**G**) k = 20. Molecular weight markers (as in Fig. 2) are on the leftmost and rightmost lanes of each gel.

capable of detecting a single template. A more realistic assumption might be that at least 5000 templates were required to produce a positive signal after just 25 cycles. With this assumption, the average probability of a correct strand surviving a single capture-release step is at least $(3.75 \times 10^{-2})^{(1/24)} = 0.87$.

Analysis for incorrect strands, those encoding truth assignments that failed to satisfy Φ, was as follows. Continue to assume that at least 5000 templates were required to produce a positive signal after 25 cycles of PCR. PCR of 1 μl of the 45 μl answer stock revealed a band corresponding to incorrect truth assignments (*25*). However, PCR of 1 μl of a 10-fold dilution of the answer stock revealed no bands corresponding to incorrect truth assignments. Hence, it is probable that at most, $10 \times 45 \times 5000 = 2{,}250{,}000$ incorrect strands were in the answer stock. It is reasonable to assume that all of these incorrect strands encoded truth assignments with a single incorrect variable assignment. Such strands are termed "1-incorrect strands." For each 1-incorrect strand, there is typically a single capture layer to which it should not anneal. There are approximately 6×10^9 1-incorrect strands at the beginning of the computation. It follows that the probability of 1-incorrect strands being accidentally retained during the critical capture-release step is at most $(2{,}250{,}000)/(6 \times 10^9) = 3.75 \times 10^{-4}$.

This analysis suggests that beginning with 500 pmol of full library, 3-SAT problems of approximately 30 variables could be solved by the means described here. Periodic PCR amplifications during the course of the computation might extend this approach to 3-SAT problems of considerably larger size.

Conclusions. For thousands of years, humans have tried to enhance their inherent computational abilities using manufactured devices. Mechanical devices such as the abacus, the adding machine, and the tabulating machine were important advances. Yet it was only with the advent of electronic devices and, in particular, the electronic computer some 60 years ago that a qualitative threshold seems to have been passed and problems of considerable difficulty could be solved. It appears that a molecular device has now been used to pass this qualitative threshold for a second time.

In our study, a minimalistic approach was taken. A 20-variable instance of a 3-SAT problem was solved using (except during input and output) DNA Watson-Crick pairing and melting as the sole "operation." Though computational theory would predict it, it is nonetheless remarkable that this basic molecular interaction could sustain such a complex computation. Our implementation was that of a reduced version of the Sticker Model (*9*). We did not implement stickers and, hence, our library strands behaved like fixed memories. With stickers as originally conceived, library strands would act as more powerful write-once memories. Recent research (*27*) suggests that DNA "strand invasion" might provide a means for the specific removal of stickers from library strands. This could give rise to library strands that act as very powerful read-write memories. Further investigation of this possibility seems worthwhile.

Despite our successes, and those of others, in the absence of technical breakthroughs, optimism regarding the creation of a molecular computer capable of competing with electronic computers on classical computational problems is not warranted. However, molecular computers can be considered in a broader context. They may be useful in specialized environments where, for example, extreme energy efficiency or extraordinary information density is required. They may provide a much-needed means for controlling chemical/biological systems in the same way that electronic computers have provided a means for controlling electrical/mechanical systems. They do provide a focus for the integration of ideas from biology and computation and this can lead to spin-offs, such as the promising work on DNA self-assembly (*11*). They enlighten us about alternatives to electronic computers and studying them may ultimately lead us to the true "computer of the future."

Most importantly, DNA computers, such as the one presented here, illustrate that biological molecules can be used for distinctly nonbiological purposes. For such purposes, these molecules represent an untapped legacy of 3 billion years of evolution, and there is great potential in their further exploration.

References and Notes

1. L. M. Adleman, *Science* **266**, 1021 (1994).
2. R. Lipton, *Science* **268**, 542 (1995).
3. E. B. Baum, *Science* **268**, 583 (1995).
4. P. W. K. Rothemund, in *DNA Computers*, R. J. Lipton, E. B. Baum, Eds., vol. 27 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* (American Mathematical Society, Providence, RI, 1996), pp. 75–119.
5. W. D. Smith, in (*4*), pp. 121–185.
6. Q. Quyang, P. D. Kaplan, S. Liu, A. Libchaber, *Science* **278**, 446 (1997).
7. T. Head, *Bull. Math. Biol.* **49**, 737 (1987).
8. F. Guarnieri, M. Fliss, C. Bancroft, *Science* **273**, 220 (1995).
9. S. Roweis *et al.*, *J. Comp. Biol.* **5**, 615 (1998).
10. L. Landweber, L. Kari, *BioSystems* **52**, 3 (1999).
11. E. Winfree, F. R. Liu, L. A. Wenzler, N. C. Seeman, *Nature* **394**, 539 (1998).
12. N. Morimoto, M. Arita, A. Suyama, in *DNA Based Computers III*, H. Rubin and D. H. Wood, Eds., vol. 48 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* (American Mathematical Society, Providence, RI, 1999), pp. 193–206.
13. K. Sakamoto *et al.*, *Science* **288**, 1223 (2000).
14. Q. Liu *et al.*, *Nature* **403**, 175 (2000).
15. C. D. Mao, T. H. LaBean, J. H. Reif, N. C. Seeman, *Nature* **407**, 493 (2000).
16. D. Faulhammer, A. R. Cukras, R. J. Lipton, L. F. Landweber, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1385 (2000).
17. Y. Benenson *et al.*, *Nature* **414**, 430 (2001).
18. H. Yoshida, A. Suyama, in *DNA Based Computers 5*, E. Winfree and D. K. Gifford, Eds., vol. 54 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* (American Mathematical Society, Providence, RI, 1999), pp. 9–20.
19. R. S. Braich *et al.*, in *Lecture Notes In Computer Science*, A. Condon and G. Rozenberg, Eds. (Springer-Verlag, New York, vol. 2054, 2000), pp. 27–42. The program for generating the sequences may be obtained at www.usc.edu/dept/molecular-science/.
20. T. H. Cormen, C. E. Leiserson, R. L. Rivest, *Introduction to Algorithms* (The MIT Press, Boston, MA, 1990), pp. 916–963.
21. M. Kenney, S. Ray, T. Boles, *BioTechniques* **25**, 516 (1998).
22. Value sequences X^T_k and X^F_k, (k = 1...20). All sequences written 5′ to 3′. X^T_1 = TTA CAC CAA TCT CTT, X^F_1 = CTC CTA CAA TTC CTA, X^T_2 = ATT TCC AAC ATA CTC, X^F_2 = AAA CCT AAT ACT CCT, X^T_3 = TCA TCC TCT AAC ATA, X^F_3 = CCC TAT TAA TCA ATC, X^T_4 = TCA CTC CAC TTA ACT, X^F_4 = TAC TTA TAA CTT CCC, X^T_5 = ATA ACC ACA AAC TCA, X^F_5 = TCT CAA TAC CAC CTA, X^T_6 = CTA TCC AAT AAC CTC, X^F_6 = TTC ATA CAC TTA CAC, X^T_7 = TTC CAC CCC AAT AAA, X^F_7 = AAC TCA TAC TAC TCA, X^T_8 = CTA TTT ATA TCC ACC, X^F_8 = TAT TCT CAC CCA TAA, X^T_9 = ACA CCT AAC TAA ACT, X^F_9 = ACA CTA TCA ACA TCA, X^T_{10} = CTA CCC TAT TCT ACT, X^F_{10} = CCT TTA CCT CAA TAA, X^T_{11} = ATC TTT AAA TAC CCC, X^F_{11} = CTC CCA AAT AAC ATT, X^T_{12} = TCC ATT TCT CCA TAT, X^F_{12} = AAC TTC ACC CCT ATA, X^T_{13} = TTT CTT CCA TCA CAT, X^F_{13} = TCA TAT CAA CTC CAC, X^T_{14} = CAT TCA ATC CAC TAC, X^F_{14} = ACC CAA TCC TCT TAA, X^T_{15} = AAC AAC CTT ATC CTT, X^F_{15} = TAA TAA CCC ATC CTA, X^T_{16} = TCA CTA CAT TAC CTT, X^F_{16} = TCA TCA AAC CTC ACA, X^T_{17} = ACA AAC CCT AAC ATT, X^F_{17} = CTC AAC AAT TTT CCA, X^T_{18} = TCT TAC CAT CTT CAT, X^F_{18} = AAC ACA TTA CTT CCT, X^T_{19} = CTC TTC TCC TCT TTT, X^F_{19} = ACC CAT TAC TAC CAT, X^T_{20} = ACA CAA ATA CAC ATC, X^F_{20} = CAA CCA AAC ATA AAC.
23. Web figure 1 is available on *Science* Online at www.sciencemag.org/cgi/content/full/1069528/DC1.
24. Standard conditions: 50 μl total volume containing 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris (pH 8.8), 200 μM of each dNTP, 10 pmol of each primer, and 1 unit of Taq DNA Polymerase, on a GeneAmp PCR System 9700 (Applied Biosystems) with the following temperature profile: 95°C for 15 s, 40°C for 45 s, 72°C for 60 s.
25. R. S. Braich *et al.*, data not shown.
26. W. P. C. Stemmer, A. Crameri, K. D. Ha, T. M. Brennan, H. L. Heyneker, *Gene* **164**, 49 (1995).
27. B. Yurke, A. Turberfield, A. P. Mills, F. C. Simmel, J. L. Neumann, *Nature* **406**, 605 (2000).
28. Modules were inserted into the electrophoresis box in pairs. The exterior of each module was sheathed with tape at the end furthest from the capture layer. A sleeve was created from the top 6 mm of a 20 μl pipette tip (RT-L10, Rainin). The modules were inserted taped end first into opposite ends of the sleeve. To ensure good contact, a 5- to 10-μl aliquot of 5% polyacrylamide solution was placed between them. The modules were manually adjusted to ensure that they touched and were in correct alignment. To accommodate modules, a 1.5-cm-diameter hole was drilled through the plexiglass divider of the electrophoresis box and fitted with a PE Quickdisconnect tube connector (for tubing of inner diameter 1/8 inches to 1/4 inches, Aldrich) from which the nozzles were cut off.

29. We gratefully acknowledge the contributions of M. Goodman, A. Soltani, D. Hwang, L. Kari, N. Chelyapov Jr., and the members of the Laboratory for Molecular Science. Supported by grants from NASA/Jet Propulsion Laboratory, the Defense Advanced Research Projects Agency, the Office of Naval Research, and NSF.

3 January 2002; accepted 5 March 2002
Published online 14 March 2002;
10.1126/science.1069528
Include this information when citing this paper.