

# Demonstration of a universal surface DNA computer

Xingping Su and Lloyd M. Smith\*

Department of Chemistry, University of Wisconsin–Madison, 1101 University Avenue, Madison, WI 53706, USA

Received April 19, 2004; Accepted May 13, 2004

## ABSTRACT

**A fundamental concept in computer science is that of the universal Turing machine, which is an abstract definition of a general purpose computer. A general purpose (universal) computer is defined as one which can compute anything that is computable. It has been shown that any computer which is able to simulate Boolean logic circuits of any complexity is such a general purpose computer. The field of DNA computing was founded in 1994 by Adleman's solution of a 7-bit instance of the Hamiltonian path problem. This work, as well as most of the subsequent experimental and theoretical investigations in the area, focused primarily upon the solution of NP-complete problems, which are a subset of the larger universal class of problems. In the present work a surface DNA computer capable of simulating Boolean logic circuits is demonstrated. This was done by constructing NOR and OR gates and combining them into a simple logic circuit. The NOR gate is one of the universal gates in Boolean logic, meaning that any other logic gate can be built from it alone. The circuit was solved using DNA-based operations, demonstrating the universal nature of this surface DNA computing model.**

## INTRODUCTION

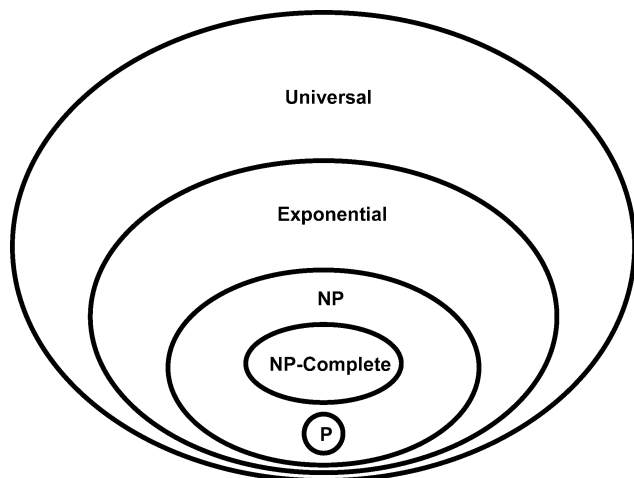
DNA computing was proposed as an alternative computing paradigm to electronic computers for solving NP-complete problems. After Adleman's seminal experiment (1), many research groups developed their DNA computing models by tackling different instances of the NP-complete problem, in particular the 3-satisfiability (3-SAT) problem, a benchmark for testing the performance of various DNA computers. Liu *et al.* used a surface chemistry based DNA computer to solve a four-variable four-clause 3-SAT problem (2). Yoshida and Suyama solved a four variable SAT problem using a DNA program implementing a breadth first search (3). Sakamoto *et al.* solved a six-variable SAT problem using hairpin DNA (4). Another group led by Landweber used RNA to solve an instance of a nine variable satisfiability problem related to the 'knight's problem' in chess (5). Recently, Adleman's team solved a 20-bit 3-SAT problem (6). This is by far the largest problem solved by any DNA computer to date.

In computational complexity theory, the class of NP-complete is a subset of the universal class (7), as shown in Figure 1. Fewer studies, however, have been reported on the universality of DNA computers. Boneh *et al.* demonstrated that DNA computers can theoretically solve computational problems beyond the NP class (8). Winfree demonstrated that universal computations are possible with self assembly-based computation, both theoretically and experimentally (9). Benenson *et al.* (10) described a programmable finite molecular automaton comprised of DNA and DNA-manipulating enzymes that solved computational problems autonomously. This automaton is a special case of the universal Turing machine. The transition molecules of this automaton, however, can only 'read' its input tape encoded in a double-stranded DNA, whereas the transition functions of the universal Turing machine can both 'read' and 'write' on its tape. This automaton, therefore, is less powerful than the universal Turing machine and is not a universal computer. Stojanovic and Stefanovic (11) constructed deoxyribozyme-based logic gates and a molecular automaton called MAYA, which encoded a version of the game tic-tac-toe and interactively competed against a human opponent. Their logic gates could conceivably be used to construct Boolean logic circuits more complex than MAYA, although the universality of such an approach was not discussed.

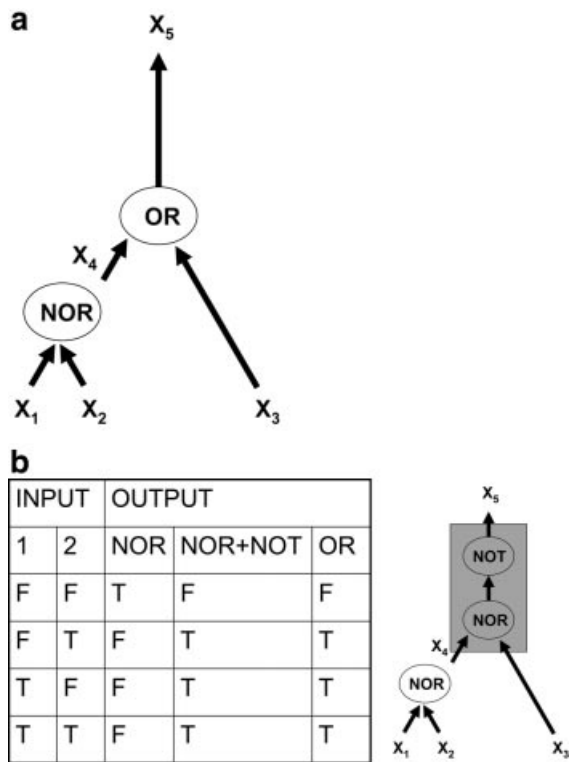
Our group developed a surface-based DNA computer to solve a four-variable four-clause 3-SAT problem (2). Briefly, 16 synthetic DNAs encoding all truth-value assignments for the four binary variables were attached to a planar gold-coated glass substrate. Successive cycles of hybridization (MARK operation), exonuclease digestion (DESTROY operation) and denaturation (UNMARK operation) were used to identify and eliminate those DNAs encoding truth-values that did not satisfy each of the four clauses of the 3-SAT problem. Upon completion of four cycles of such operations, the DNAs remaining on the surface were those that satisfied the 3-SAT problem. These molecules were then identified in a READOUT operation. In related work, our group developed a multiple-word encoding strategy (12,13), which was employed here. In addition, a theoretical approach based on the surface DNA computing model for computing Boolean logic circuits has been reported (14).

In this study, the surface DNA computing model is shown to be capable of simulating a Boolean logic circuit. The NOR gate is a 'universal gate' in Boolean logic (15), meaning that any other logic gate can be built from it. In this proof of principle experiment, it is shown that a NOR gate can be built based on the surface computing model, that an OR gate can be built from such a NOR gate, and that such a NOR gate and an

\*To whom correspondence should be addressed. Tel: +1 608 263 2594; Fax: +1 608 265 6780; Email: smith@chem.wisc.edu

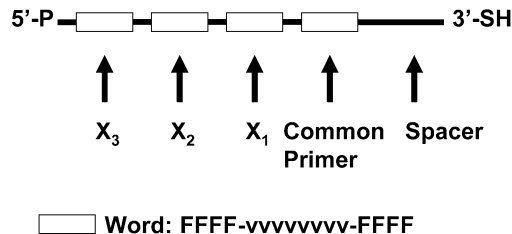


**Figure 1.** Computational complexity classes. Note that although the class P and NP-Complete are widely believed to be disjoint, as depicted here, there is no proof of this.



**Figure 2.** (a) A circuit consisting of a NOR gate and an OR gate. (b) Truth-value table for an OR gate, a NOR gate and a combination of a NOR and a NOT gate, and an equivalent circuit to the one in (a).

OR gate can be combined to construct a circuit. For simplicity, this circuit had only two inputs to each gate (Fig. 2a) and had three initial input variables. Each variable was encoded in two 16 nt DNA sequences (called words), one for TRUE and the other for FALSE. Three such words, one for each variable, were strung together to encode one truth-value combination for the three variables; thus, there were eight input DNAs. The



**Figure 3.** Sequences (5'→3') of the DNA encoding the inputs. Three 16 nt words encoded three bits of information:  $X_1$ ,  $X_2$  and  $X_3$ . They were followed by a common primer site for polymerase extension. Each word was made up of an 8 nt encoding region, v8, and a 4 nt word fixed label, F4, at either end. Ten phosphoramidite-18 spacer units were inserted between the nucleotides and the thiol group to facilitate hybridization (13). These DNA oligonucleotides were attached to the surfaces through the thiol group at their 3' end.

**Table 1.** Sequences of the words (5'→3')

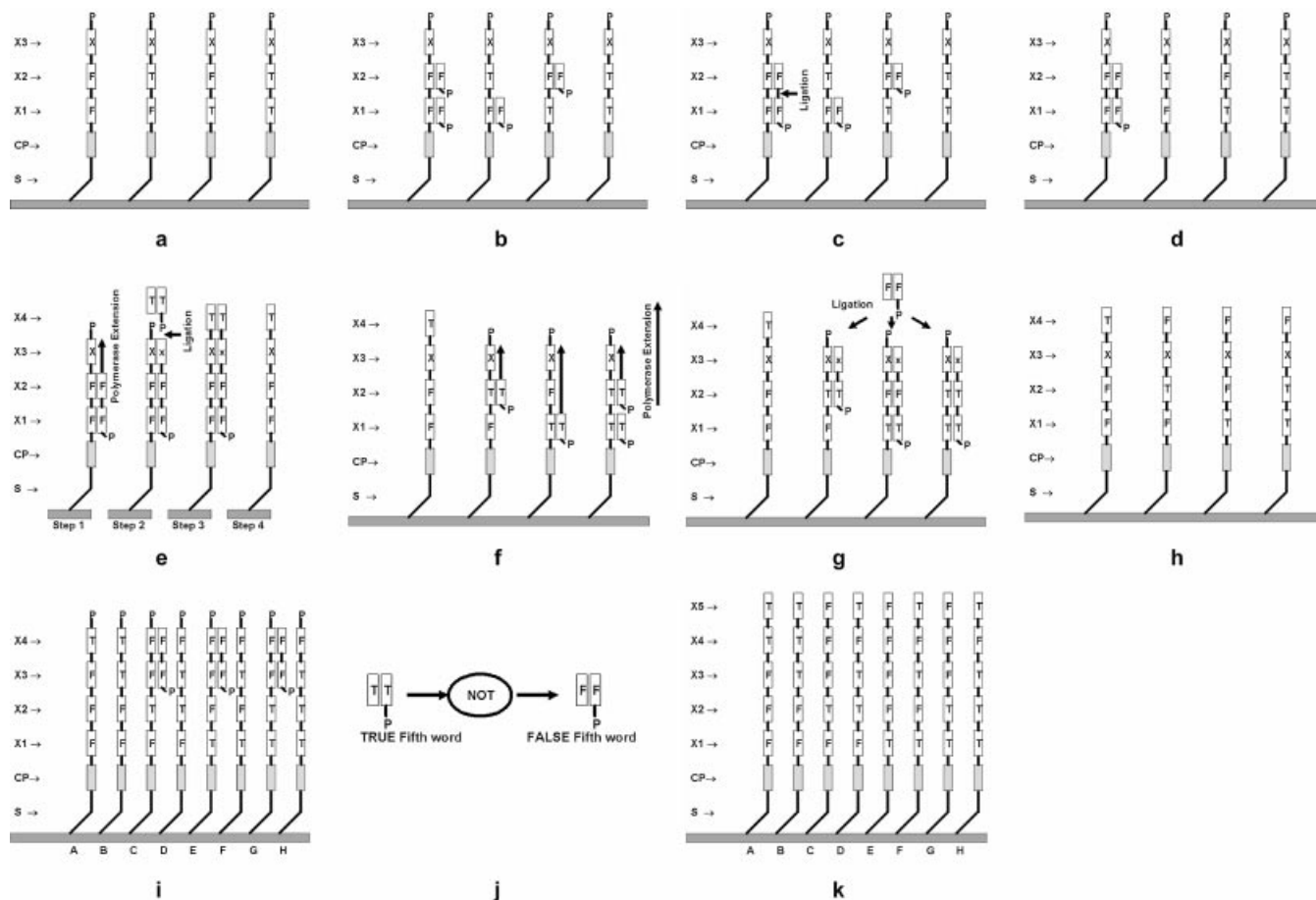
	TRUE	FALSE
$X_1$	AACGttcctgcaGCAA	AACGaaccaccaGCAA
$X_2$	TTGCatcgagctCGTT	TTGCtccttacgCGTT
$X_3$	CGAAAtacgtctAAGC	CGAAccaacaacAAGC
$X_4$	CGTTTcगतageTTGC	CGTTtaggaatgcTTGC
$X_5$	TTAAAtatcgcgAATT	TTAAAtaatgccgAATT

**Table 2.** Sequences of the word complements (5'→3')

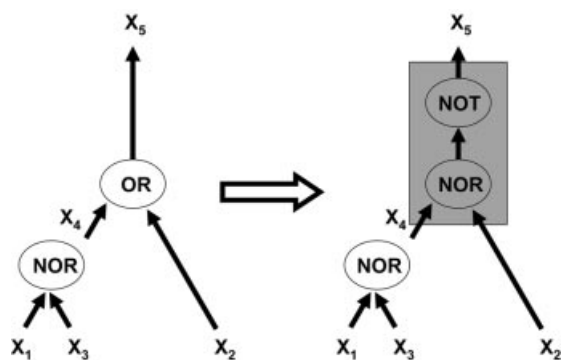
	TRUE	FALSE	LNA/DNA (FALSE)
$X_1$	TTGCtgcaggaaCGTT	TTGCtgggtggtCGTT	t+gg+tg+gt+t
$X_2$	AAGCagctcgatGCAA	AAGCgtaaggaGCAA	c+gt+aa+gg+a
$X_3$	GCTTtaggacgtaTTCG	GCTTgttgggtTTCG	g+tt+gt+tg+g
$X_4$	GCAAgtatcgaAACG	GCAAgtccttAACG	g+ca+tt+cc+t
$X_5$	AATTcgcgatatTTAA	AATTcggcattaTTAA	

sequences of these input DNAs are shown in Figure 3. The sequences of the constituting words are listed in Table 1 and those of the complements to the words in Table 2. The input DNAs were attached to the gold slide surfaces in an 'addressed' fashion, i.e. it was known where each DNA was attached (Figs 4a and 7b). To compute the circuit, first a NOR gate was built to compute ( $X_1$  NOR  $X_2$ ). A fourth word,  $X_4$ , registering the results from the NOR gate, was appended to the 5' end of the attached input DNAs. Then another NOR gate and a NOT gate were combined to build an OR gate to compute ( $X_3$  OR  $X_4$ ). The results from this OR gate were registered by a fifth word,  $X_5$ , that was appended to the input DNAs.

The NOR gate of the circuit was built in two rounds of MARK and APPEND-MARKED operations. A MARK operation marks, i.e. identifies, DNAs encoding certain input variables and an APPEND-MARKED operation appends a new word to such MARKED DNAs. The molecular steps employed to effect these operations are described below. The first round of such operations marked the input DNAs in which

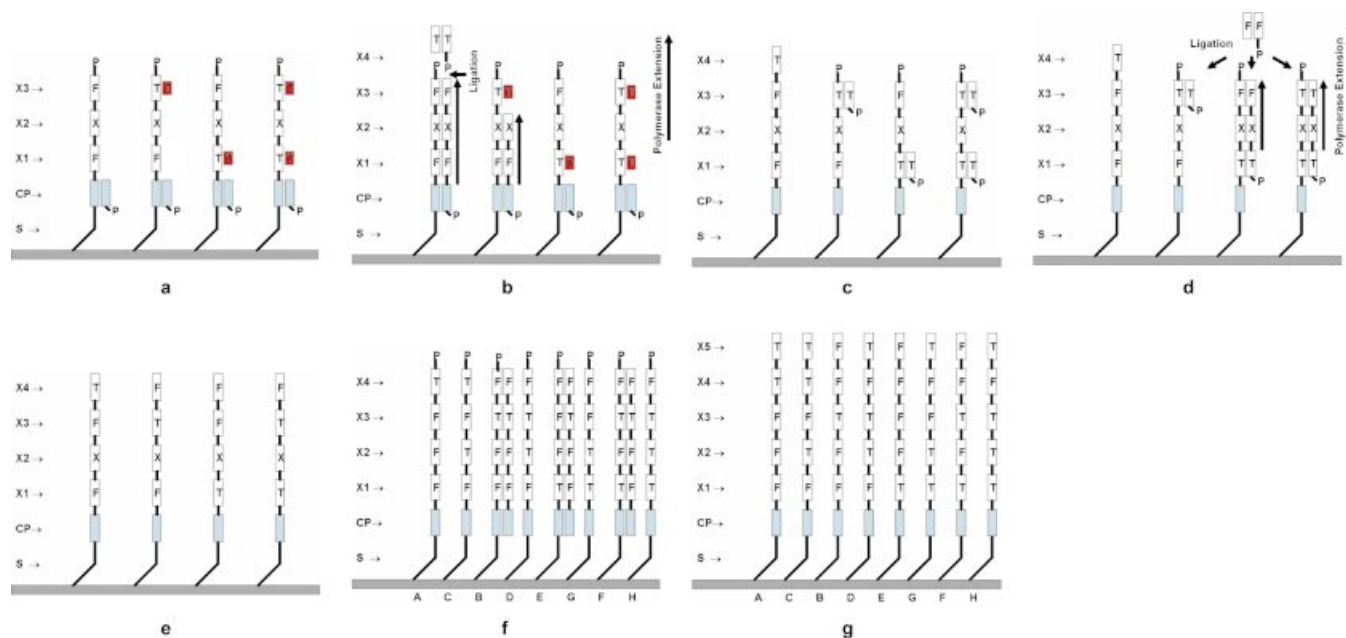


**Figure 4.** Computation of a simple circuit consisting of a NOR gate and an OR gate. (a) DNAs encoding the inputs were attached to the surfaces by the 3' thiol group. 'FFX' designates the input DNAs in which  $X_1$  is FALSE,  $X_2$  is FALSE and  $X_3$  can be either TRUE or FALSE (i.e. input A and B). P indicates a phosphate group and S the underlying surface chemistry. (b) Complements to the words ( $X_1$  is F) and ( $X_2$  is F) were added to the surfaces and allowed to hybridize with the attached DNAs. (c) T4 DNA ligase linked the two complements hybridized with the 'FFX' DNAs. (d) Differential melting melted off hybridized one-word complements and left the linked two-word complements in place. (e) In the 'FFX' DNAs, polymerase extended the two-word complements to their 3' ends. A TRUE fourth word ( $X_4$  is T) was appended to the distal end of the now double-stranded 'FFX' DNAs. Note here that the 5' end of the newly attached word was not phosphorylated, whereas the 5' end of its complement was phosphorylated. This was designed so that only one copy of this word would be appended. (f) The complements to the words ( $X_1$  is T) and ( $X_2$  is T) were added to the surface and allowed to hybridize. (g) The complements added in (f) were extended by polymerase. Then a FALSE fourth word was appended to the surfaces by ligation. Note again that the design of this one-bit word was such that only the 5' end of its complementary strand was phosphorylated for the same reason as for (e). The lack of a 5' phosphate at the ends of the 'FFX' DNAs also ensured that this FALSE fourth word would not be appended to them. (h) DNAs after the NOR gate computation. Note here that the DNAs were now without a phosphate group. A kinase phosphorylated them before the OR gate computation. (i) DNAs in which ( $X_3$  is F) and ( $X_4$  is F) were MARKED in the first APPEND-MARKED operation for the OR gate. (j) A FALSE fifth word, rather than a TRUE fifth word, was appended. The change is equivalent to a NOT gate. (k) DNAs after the OR gate computation.



**Figure 5.** A circuit consisting of a NOR gate and an OR gate, when the two input words are non-contiguous

both input variables to the NOR gate are FALSE, i.e. both ( $X_1$  is F) and ( $X_2$  is F), and appended a TRUE fourth word to them. The second round identified the other input DNAs, i.e. those in which [( $X_1$  is F) and ( $X_2$  is T)], [( $X_1$  is T) and ( $X_2$  is F)] or [( $X_1$  is T) and ( $X_2$  is T)] and appended a FALSE fourth word to them. In the first MARK operation, the complements to the words encoding ( $X_1$  is F) and ( $X_2$  is F) were added to the surfaces and allowed to hybridize with the attached input DNAs (Fig. 4b). Both complements were hybridized only with those DNAs to be marked, i.e. those designated 'FFX'. They were linked by T4 DNA ligase (Fig. 4c), forming a two-word complement. A differential melting procedure removed the one-word hybridized complements while leaving the two-word complement in place (Fig. 4d). This two-word extension then served as a primer for a polymerase extension



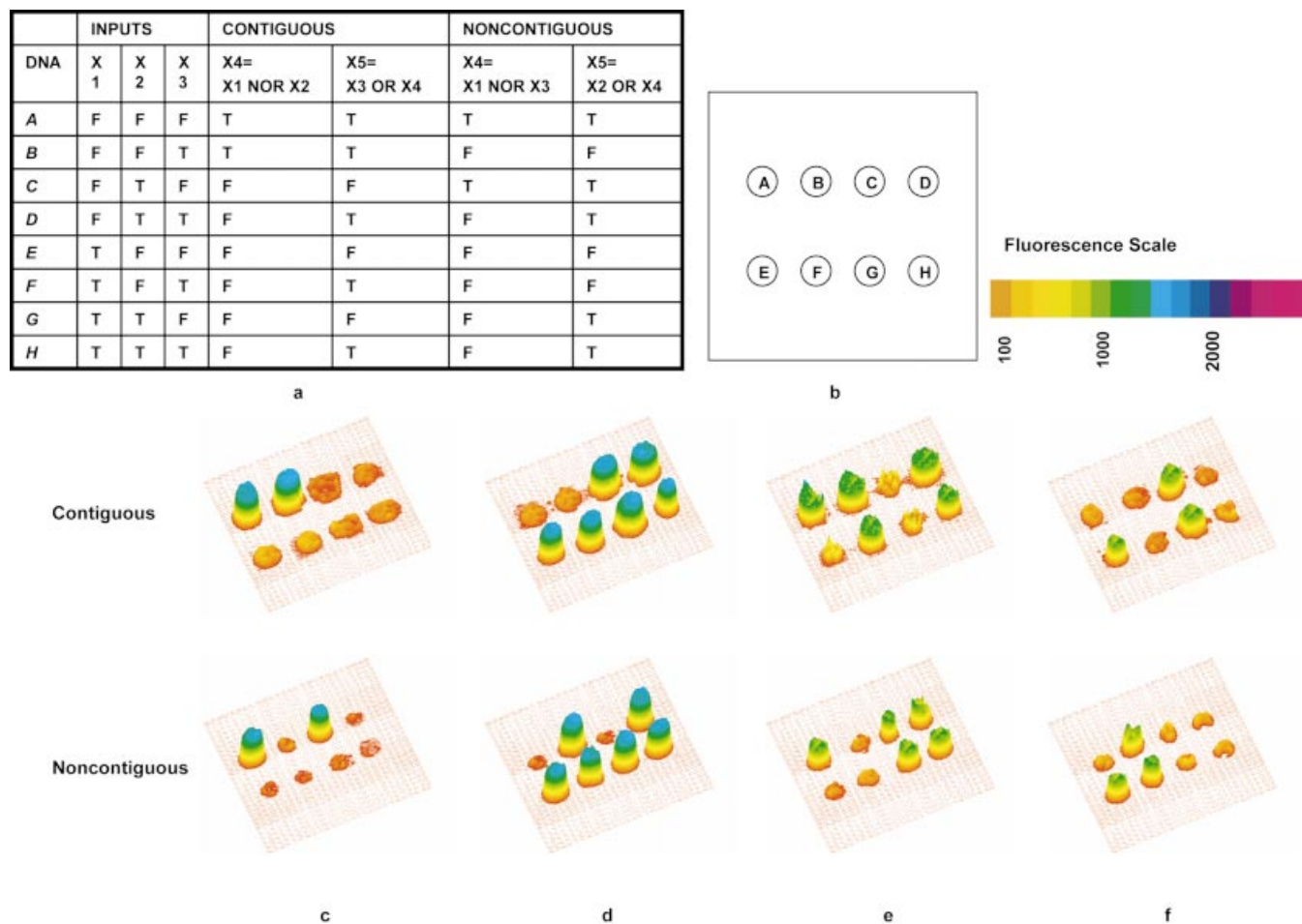
**Figure 6.** Computation of a simple circuit consisting of a NOR gate and an OR gate, when the two inputs to a gate are non-contiguous. (a) The complements (red) to the words ( $X_1$  is T) and ( $X_3$  is T), both chimeras of LNA/DNA, were added to the surface and allowed to hybridize. The complement (light green) to the common primer was also hybridized. Note also the difference in length. Chimeras were shorter 8mers. (b) Polymerase extended the 'FXF' DNAs to the distal end, allowing a TRUE fourth word to be appended to them. In other DNA oligonucleotides, this TRUE fourth word was not appended. (c) Complements to the words ( $X_1$  is T) and ( $X_3$  is T) of regular DNA were then hybridized to the surfaces. (d) These hybridized complements served as primers for polymerase extension, resulting in a double-stranded structure at the distal ends of the attached DNAs for appending a FALSE fourth word. (e) Input DNAs after the NOR gate computation. (f) The first MARK operation for the OR gate. (g) Input DNAs after the OR gate computation.

reaction, which resulted in a double-stranded structure in the MARKED DNAs (Fig. 4e), allowing a TRUE fourth word to be appended to them. The extended complements to the 'FFX' DNAs were then melted off. In the second MARK operation (Fig. 4f), the complements to the words encoding ( $X_1$  is T) and ( $X_2$  is T) were added to the surfaces. They hybridized to all attached DNAs but the 'FFX' ones and served as primers for another polymerase extension reaction that resulted in a double-stranded structure of the MARKED DNAs. A FALSE fourth word was then appended to them (Fig. 4g). To validate the NOR gate computation (Fig. 4h), the fluorescently tagged (FT) complement to the newly appended TRUE fourth word and that to the FALSE one were added to the surfaces in separate experiments. The chips were scanned in a fluorescence imager. The presence or absence of fluorescence signals from the spots where DNAs were attached revealed the identities of the appended words.

The truth-value table in Figure 2b shows that an OR gate is equivalent to the combination of a NOR gate and a NOT gate. Consequently, the OR gate in the circuit was constructed similarly to the NOR gate. The same MARK operations marked DNAs and the same APPEND-MARKED operations appended a fifth word, registering the OR gate computation. In the first MARK operation, the DNAs in which both ( $X_3$  is F) and ( $X_4$  is F) were marked (Fig. 4i). In the first APPEND-MARKED operation of the above NOR gate computation, a TRUE fourth word was appended to such DNAs. In the case of the OR gate, however, a FALSE fifth word was appended. This change was equivalent to a NOT gate computation (Fig. 4j). Similarly, in the second APPEND-MARKED

operation of the OR gate computation, a TRUE fifth word was appended to the other DNAs after they were marked in the second MARK operation.

In the above computation, the two words encoding the two input variables to a gate were next to each other, i.e. contiguous. When such words were not contiguous, for example when the same input DNAs used above computed the circuit in Figure 5, a totally different MARK operation was required. In the first MARK operation of the NOR gate in the above contiguous case, complements to the words ( $X_1$  is F) and ( $X_3$  is F) were hybridized to the surface. In the non-contiguous case, however, complements (8 nt in length) to the words ( $X_1$  is T) and ( $X_3$  is T) were hybridized to the surfaces. These complements were chimeras of locked nucleic acids (LNAs) (16) and regular DNA. Their sequences are listed in Table 2. LNAs are a type of modified nucleic acid. When hybridized to the same complements, LNA/DNA chimeras have higher melting temperatures than do their regular DNA counterparts of the same sequence. Here such LNA/DNA chimeras were used to block the polymerase extension reaction. Previously, peptide nucleic acids (PNAs) have been used for the same purpose (13). One significant limitation of PNAs is that the synthesis of PNAs is less general than that of normal DNAs, whereas this is not an issue for the synthesis of LNAs. The sequences of alternating DNA and LNA of the LNA/DNA chimeras (16) were designed to give a high degree of polymerase blocking. After the complements were allowed to hybridize, those input DNAs to be MARKED hybridized none of the chimera complements, whereas the other DNAs had either one or both complements hybridized to them. Then



**Figure 7.** Results from the circuit computations. (The program NIH Image, freely available at <http://rsb.info.nih.gov/nih-image/>, was used to process these fluorescence images for presentation.) (a) The truth-table for the circuit shows the expected results to input DNAs A–H in both the contiguous and non-contiguous cases. (b) DNA oligonucleotides A–H were attached to the gold slide surfaces in an ‘addressed’ fashion. (c) After the NOR gate, the FT complement to the TRUE fourth word was hybridized to the surface and the slide was scanned. (d) After the NOR gate, the FT complement to the FALSE fourth word was hybridized to the surface and the slide was scanned. (e) After the OR gate, the FT complement to the TRUE fifth word was hybridized to the surface and the slide was scanned. (f) After the OR gate, the FT complement to the FALSE fifth was hybridized to the surface and the slide was scanned.

the complement to the common primer was added (Fig. 6a). In the ‘FXF’ DNAs, the ones to be MARKED, polymerase extended the primer complement all the way up to the distal end, forming a double-stranded structure for appending a fourth TRUE word. In the other strands, the LNA/DNA complements not only blocked the polymerase extension starting from the common primer complement but also failed to serve as a primer for polymerase extension from their 3’ end, thus preventing the formation of a blunt-end for the subsequent ligation. Consequently, the TRUE fourth word was appended to ‘FXF’ DNAs alone (Fig. 6b). These LNA/DNA chimera complements were then melted off. For the second MARK operation for this NOR gate, the complements to the words ( $X_1$  is T) and ( $X_3$  is T), both regular DNAs, were hybridized to the surfaces (Fig. 6c) and served as primers for another polymerase extension, which resulted in a double-stranded structure of the DNAs to be marked (Fig. 6d). A FALSE fourth word was appended to these MARKED DNAs (Fig. 6e).

In the OR gate computation, these new MARK operations identified the DNAs in which both ( $X_2$  is F) and ( $X_4$  is F)

(Fig. 6f) and a FALSE fifth word was appended to them. Then the other DNAs were marked and a TRUE fifth word appended to them similarly (Fig. 6g).

## MATERIALS AND METHODS

### Materials

The chemicals 11-amino-1-undecanethiol hydrochloride (MUAM) (Dojindo, Gaithersburg, MD), sulfosuccinimidyl 4-[*N*-maleimidomethyl]-cyclohexane-1-carboxylate (SSMCC) (Pierce, Rockford, IL), 40% acrylamide/bis-acrylamide solution, ammonium persulfate, TEMED, urea (Bio-Rad Laboratories, Hercules, CA) and triethanolamine hydrochloride (TEA) (Sigma, Milwaukee, WI) were all used as received. Another chemical, 18-mercapto-octadecyl amine,  $\text{HS}(\text{CH}_2)_{18}\text{NH}_2$ , was synthesized in-house by Dr D. Peelen (manuscript in preparation). Millipore filtered water was used for all aqueous solutions. DNAs were synthesized on an ABI DNA synthesizer at the Biotechnology Center at the University of Wisconsin or obtained from Integrated DNA

technologies (Coralville, IA). Chemical Phosphorylation Reagent for 5'-phosphorylation, 3'-thiol-modifier C3 S-S CPG for 3'-thiol modification, 6-FAM for fluorescein labeling and spacer phosphoramidite-18 were from Glen Research (Sterling, VA). LNA/DNA chimeras were obtained from Prologo (Boulder, CO). Deep Vent (*exo*<sup>-</sup>) DNA polymerase, T4 polynucleotide kinase, T4 DNA ligase and their supporting materials were from New England BioLabs (NEB) (Beverly, MA).

### Sequence design of DNAs and DNA/LNA chimeras

The design tools at the Prologo website (<http://www.prologo.com>) were used to select the word sequences from the 108 8mers reported in Frutos *et al.* (12), so as to minimize the secondary structures that could form in their LNA/DNA chimeras. The sequences of the words are listed in Table 1. There were three sets of complements to the words. The first set included complements to X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub> and X<sub>4</sub>. They were chemically phosphorylated at the 5' end when synthesized. The second set included FT complements to X<sub>3</sub>, X<sub>4</sub> and X<sub>5</sub>. The third set included four LNA/DNA complements to four FALSE words. The sequences of the regular DNA (5'-phosphate and fluorescein not indicated) and LNA/DNA were listed in Table 2. The '+' sign indicates that the nucleotide after it is a LNA.

### DNA attachment

Before attachment, thiol-modified DNAs were deprotected as outlined in the Glen Research Corp. catalog and purified by reverse phase binary gradient elution HPLC (Shimadzu SCL-6A). Their concentrations were verified prior to use by UV absorption measurements with an HP8452A UV-VIS spectrometer. Generally, immediately after purification, thiol-modified DNAs were attached to gold-coated slides (Evaporated Metal Films, New York, NY) by procedures described elsewhere (17,18). Briefly, the surfaces of gold-coated slides were modified with a self-assembled monolayer (SAM) of HS(CH<sub>2</sub>)<sub>11</sub>NH<sub>2</sub> (in the experiments measuring the efficiencies of the enzymes) or HS(CH<sub>2</sub>)<sub>18</sub>NH<sub>2</sub> (in all other experiments). Then the amino group of the SAM was reacted with the heterobifunctional linker SSMCC, which created a thiol-reactive maleimide-terminated surface. The 3'-thiol oligonucleotides were then covalently attached to this surface. A 0.5 µl drop of a solution containing 0.8 mM thiol DNA in 100 mM TEA buffer, pH 7.0 was spotted onto the surface. The attachment reaction was allowed to proceed for at least 12 h in a humid chamber. After attachment, the surface was rinsed with water and soaked for at least 1 h in 2× SSPE, 0.2% SDS buffer at 37°C. The gold slides were stored in 2× SSPE, 0.2% SDS buffer at room temperature until use.

### Melting analysis

DNA melting curves were obtained by monitoring the absorbance of DNA solutions at 260 nm as a function of temperature in the above HP8452 UV-VIS spectrometer. Melting temperatures were measured in pH 7 buffer solutions containing 10 mM sodium phosphate, 1 mM EDTA, 1 mM sodium chloride and 2 µM DNA and its complement. A ramp rate of 1°C/min with a hold time of 1 min was employed over the range 25–90°C to record the melting curves. The *T<sub>m</sub>* was

determined as the temperature at which the first derivative of the UV absorbance curve was a maximum (12).

The above melting temperatures measured in solution were taken as references in the differential melting experiments on surfaces. It has been observed that when DNA oligonucleotides are attached to surfaces their *T<sub>m</sub>* values are generally decreased (19). In designing the experiments reported here, the surface *T<sub>m</sub>* was estimated to be ~5°C lower than the solution *T<sub>m</sub>*.

### Hybridizations on surfaces

A 25 µl aliquot of a solution containing 2 µM each desired complement in 2× SSPE, 0.2% SDS buffer was applied to a gold slide and spread over the entire surface and a coverslip placed on top of the solution. Hybridization was allowed to proceed for 30 min at room temperature.

### Ligation reactions

A cohesive ligation reaction was used to link two complements hybridized contiguously to the same DNA template attached on the surface (e.g. Fig. 4c). After the two complements were allowed to hybridize, the slide was rinsed with 2× SSPE, 0.2% SDS buffer. A 30 µl aliquot of a solution containing 1 U/µl T4 DNA ligase in 1× T4 DNA ligase buffer (both from NEB) was applied to the gold slide and the ligation was allowed to proceed for 3 h at room temperature.

A blunt-end ligation reaction was used to append a new word to MARKED DNAs, which were double-stranded (e.g. Fig. 4e). A solution containing 50 µM word to be appended and 50 µM phosphorylated complement was heated in a water bath to 90°C and allowed to cool down slowly. This solution was then used to prepare a ligation solution containing 5 U/µl T4 DNA ligase and 20 µM double-strand word in 1× T4 DNA ligase buffer. A 30 µl aliquot of this ligation solution was applied to the gold slide and the ligation was allowed to proceed overnight at room temperature.

### Polymerase extension

This reaction generated a double-stranded structure in the DNAs to be marked, allowing for a blunt-end ligation to append a new word. A 100 µl aliquot of a solution containing 0.1 U/µl Deep Vent (*exo*<sup>-</sup>) DNA polymerase, 200 µM dNTP and 100 µg/ml BSA in 1× ThermoPol reaction buffer (all from NEB) was applied to the gold slide and the reaction was allowed to proceed for 2 h at 37°C.

### Kinase phosphorylation

This reaction adds a phosphate group to the 5' end of the attached DNA oligonucleotides before the OR gate computation. A 100 µl aliquot of a solution containing 0.2 U/µl T4 polynucleotide kinase, 1× reaction buffer and 1 mM ATP was applied to the gold slide and the reaction was allowed to proceed for 2 h at 37°C.

### Efficiency of polymerase

A FT primer was used in the polymerase extension reaction. After the reaction, extended primer and unextended primer were collected as described (11). The mixture was then loaded on a 15% denaturing polyacrylamide gel containing 7 M urea and separated by electrophoresis. The gel was scanned with a FluorImager575 (Molecular Dynamics, Sunnyvale, CA). The

efficiency of the polymerase extension was defined as the ratio of the fluorescence intensity of the band on the gel corresponding to the extended primer over the sum of the fluorescence intensities of both the extended and the unextended bands.

LNA/DNA chimeras were employed to block polymerase extension in some of the computations. To test their blocking ability, an 8 nt chimera complementary to the second word (the variable region of the second word) of a three-word DNA attached to a gold slide, together with a 16 nt FT complement to a common primer site in the same three-word DNA, were hybridized to the surfaces. In one control experiment, the chimera was replaced with its regular DNA counterpart of the same sequence; in another, there was no blocking agent added to the surface. Then a polymerase reaction was carried out on the gold slides. After the reaction, the extended complements to the primers were melted off and collected and analyzed by electrophoresis in separate lanes on a 15% denaturing polyacrylamide gel containing 7 M urea. The gel was then scanned by the above FluorImager.

### Efficiency of ligase

To measure the efficiency of the cohesive ligation reaction, an FT complement to the second word and an untagged one to the third word were hybridized to a three-word DNA attached to a substrate as described and then the two complements were linked by a ligase. The linked two-word complement and any unlinked FT complement to the second word were melted off from the surfaces. The mixture was then separated by electrophoresis in a 15% denaturing polyacrylamide gel containing 7 M urea. The gel was then scanned with the FluorImager. The efficiency of the cohesive ligation reaction was defined as the ratio of the fluorescence intensity of the band corresponding to the two-word complement to the sum of the fluorescence intensities of the two-word complement band and the one-word complement band.

The efficiency of the blunt-end ligation was measured differently. A three-word DNA was attached to two gold slide surfaces. An FT complement to the third word was hybridized to both slides. A fourth word was then appended to one of the slides in a blunt-end ligation reaction, forming a two-word FT complement. On the other gold slide there was no ligation reaction and the FT complement remained one-word long. Then the two-word and the one-word FT complements were melted-off and analyzed by gel electrophoresis and fluorescence measurement. The efficiency of the blunt-end ligation was defined as the ratio of the fluorescence intensity of the two-word complement to that of the one-word complement.

### Validation of the appended words

After each Boolean gate computed its inputs, the newly appended words registering the results were validated by a simple hybridization. FT complements to the newly appended TRUE and FALSE words were hybridized in separate experiments. The gold slide was scanned with the FluorImager. The results are shown in Figure 7.

### Efficiencies of computation

The efficiency of each gate on each input DNA was measured individually. After the computations, the gold slide on which the non-contiguous case was computed was cut into eight

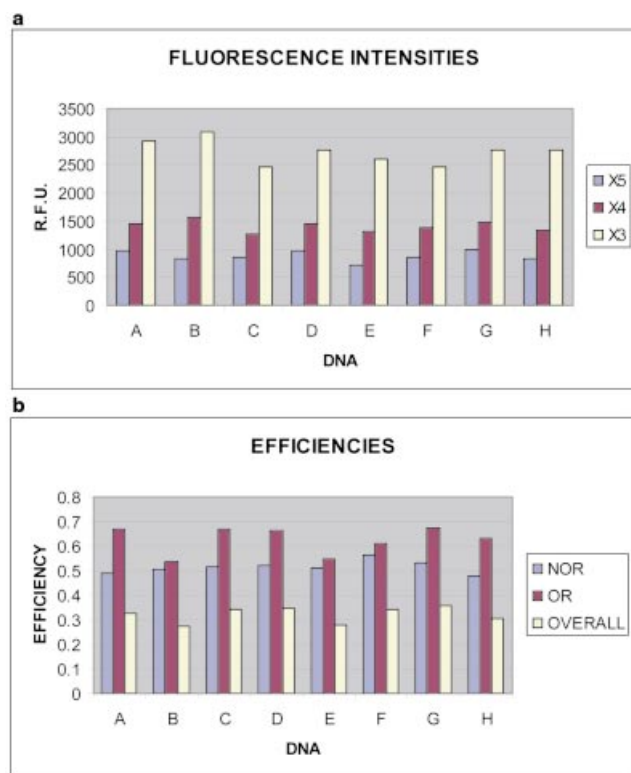
pieces, each of which had one input DNA attached. Then the FT complements to the third word, the fourth word and the fifth word were hybridized to each piece separately. Each of the fluorescent complements was then melted off, collected and loaded in a lane in a 15% denaturing polyacrylamide gel containing 7 M urea and analyzed by electrophoresis and fluorescence imaging. For a particular input DNA, the efficiency of the NOR gate computation was defined as the ratio of the fluorescence intensity of the FT complement to the correctly appended fourth word to that of the FT complement to the third word. Similarly, the efficiency of the OR gate computation was defined as the ratio of the fluorescence intensity of the FT complement to the correctly appended fifth word to that of the FT complement to the correctly appended fourth word. The overall computation efficiency was defined as the product of the efficiencies of the NOR gate and OR gate.

## RESULTS AND DISCUSSION

In this study, a Boolean logic NOR gate was built using the surface computing paradigm. The results from a NOR gate computation in Figure 7c and d clearly show that a TRUE fourth word was appended to the input DNAs in which both the two input words were FALSE, namely A and B in the contiguous case and A and C in the non-contiguous case, respectively, and that a FALSE fourth word was appended to the other input DNAs. These results are in concordance with the truth-value table. The NOR gate is combined with a NOT gate to build an OR gate and these two gates are combined into a circuit. Figure 7e and f shows that the experimental results from such circuit computations are in concordance with the truth-value table. In these experiments, the parallelism of the surface DNA computing model is maintained in that the eight input DNAs were processed simultaneously.

LNA/DNA chimeras were used in this study to block polymerase activity. In the tests to determine its blocking ability, a small amount of product was detected that resulted from a polymerase reaction that had displaced the blocking chimera (3% relative to the product from the non-displacement extension). No product was detected that would have resulted from a polymerase reaction that had occurred using the blocking chimera as a primer (data not shown). The chimeras also proved to be satisfactory in the actual computations. A recent study (20) indicates that when only one LNA nucleotide is present in a primer, there is a strong positional preference in the interaction of the LNA-modified primer and DNA polymerase. When the LNA nucleotide is at the 3' end in the primer, it has a modest effect upon polymerase extension. When it is at the penultimate position to the 3' end in the primer, it virtually completely blocks polymerase activity. The MARK and APPEND-MARKED operations using LNA/DNA chimeras in the non-contiguous case can also be used in the contiguous case without any modification.

The NOR gate is one of the universal gates in Boolean logic; if we can build a NOR gate, we can build any other gate from it alone. The construction of the OR gate serves two purposes. Firstly, as the OR gate was experimentally implemented virtually identically to the NOR gate, its construction and ensuing combination with the NOR gate into a circuit demonstrates that multiple NOR gates can be combined



**Figure 8.** Results from the overall computational efficiency experiments. (a) Background-corrected fluorescence intensities of the complements to the third word and the correctly appended fourth and fifth words. (b) Efficiencies of the NOR gate, OR gate and the overall computation.

together. This demonstration is necessary, as when other logic gates are built from the NOR gate alone, they are built by combining multiple NOR gates (15). Secondly, because in this study TRUE and FALSE are encoded in two different words, the construction of the OR gate also demonstrates that the biochemical reactions for building a NOR gate need to be only slightly adapted to build other gates. It was shown in the actual experiments that from the NOR gate to the OR gate, only a small change in the appended words was required. Other logic gates can be built in a similar fashion. For example, to build an AND gate, it suffices to MARK the DNAs in which both input variables are TRUE and APPEND a TRUE word to such MARKED DNAs and to MARK other DNAs and to APPEND a FALSE word to them. A NAND gate could also be built similarly. Thus, a circuit of any complexity can conceivably be constructed by this approach. The surface DNA computer is thus in principle capable of simulating a circuit of any complexity and is, therefore, a universal computer.

There is, of course, a considerable difference between a proof of principle experiment and a useful device. For the latter, many additional and critical considerations come into play. The most critical one is the imperfect nature of the biochemical tools employed. The efficiency of the polymerase extension was measured at 89%, that of cohesive ligation at 78% and that of the blunt-end ligation at 74% (data not shown). These results are comparable to previous studies (12,13). The imperfection of these enzymatic reactions and other biochemical processes, such as hybridizations and

melting off experiments, and the stability of the underlying surface chemistry under various experimental conditions are factors that contribute to the low overall computation efficiencies (Fig. 8). These results suggest that the number of gates that a circuit may have is quite limited at the current stage of development of surface DNA computing.

In conclusion, we have demonstrated that surface DNA computer can simulate a simple Boolean circuit and, in principle, can simulate circuits of any complexity. This proves that it is a universal computer.

## ACKNOWLEDGEMENTS

We would like to thank Dr Anne E. Condon, Dr Michael R. Shortreed and Seo Bong Chang for helpful discussions and critical comments on this manuscript. This material is based upon work supported by the National Science Foundation under Grant No. 0203892 and Grant No. 0130108.

## REFERENCES

- Adleman, L.M. (1994) Molecular computation of solutions to combinatorial problems. *Science*, **266**, 1021–1024.
- Liu, Q., Wang, L., Frutos, A., Condon, A.E. and Smith, L.M. (2000) DNA computing on surfaces. *Nature*, **413**, 175–179.
- Yoshida, H. and Suyama, A. (2000) *DIMACS: Series in Discrete Mathematics and Theoretical Computer Science. Solution to 3-SAT by Breadth First Search*. American Mathematical Society, Providence, RI, Vol. 54, pp. 9–20.
- Sakamoto, K., Gouzu, H., Komiyama, K., Kiga, D., Yokoyama, S., Yokomori, T. and Hagiya, M. (2000) Molecular computation by DNA hairpin formation. *Science*, **288**, 1223–1226.
- Faulhammer, D., Cukras, A.R., Lipton, R.J. and Landweber, L.F. (2000) Molecular computation: RNA solutions to chess problems. *Proc. Natl Acad. Sci. USA*, **97**, 1385–1389.
- Braich, R.S., Chelyapov, N., Johnson, C., Rothmund, P.W.K. and Adleman, L.M. (2002) Solution of a 20-variable 3-SAT problem on a DNA computer. *Science*, **296**, 499–502.
- Gifford, D.K. (1994) On the path to computation with DNA. *Science*, **266**, 993–994.
- Boneh, D., Dunworth, C., Lipton, R.J. and Sgall, J. (1996) On the computational power of DNA. *Discrete Appl. Math.*, **71**, 79–94.
- Winfree, E. (2000) Algorithmic self-assembly of DNA: theoretical motivations and 2D assembly experiments. *J. Biomol. Struct. Dyn.*, **11**, 263–270.
- Benenson, Y., Paz-Elizur, T., Adar, R., Keinan, Z.L. and Shapiro, E. (2001) Programmable and autonomous computing machine made of biomolecules. *Nature*, **414**, 430–434.
- Stojanovic, M. and Stefanovic, D. (2003) A deoxyribozyme-based molecular automaton. *Nat. Biotechnol.*, **21**, 1069–1074.
- Frutos, A.G., Liu, Q., Thiel, A.J., Sanner, A.M.W., Condon, A.E., Smith, L.M. and Corn, R.M. (1997) Demonstration of a word design strategy for DNA computing on surfaces. *Nucleic Acids Res.*, **25**, 4748–4757.
- Wang, L., Liu, Q., Corn, R.M., Condon, A.E. and Smith, L.M. (2000) Multiple word DNA computing on surfaces. *J. Am. Chem. Soc.*, **122**, 7435–7440.
- Cai, W., Condon, A.E., Corn, R.M., Glaser, E., Fei, Z., Frutos, T., Guo, Z., Lagally, M.G., Liu, Q., Smith, L.M. and Thiel, A. (1997) The power of surface-based DNA computation. Paper presented at the *Proceedings of The First Annual International Conference on Computational Molecular Biology (Recomb 97)*, Sante Fe, NM, January, pp. 67–74.
- Ralston, A., Reilly, E.D. and Hemmendinger, D. (2000) *Encyclopedia of Computer Science*, 4th Edn, Nature Publishing Group, London, UK.
- Petersen, M. and Wengel, J. (2003) LNA: a versatile tool for therapeutics and genomics. *Trends Biotechnol.*, **21**, 74–81.
- Brockman, J.M., Frutos, A.G. and Corn, R.M. (1999) A multi-step chemical modification procedure to create DNA arrays on gold surfaces for the study of protein-DNA interaction with surface plasmon resonance imaging. *J. Am. Chem. Soc.*, **121**, 8044–8051.

18. Chen, Y., Shortreed, M.R., Peelen, D., Lu, M and Smith, L.M. (2004) Surface amplification of invasive cleavage products. *J. Am. Chem. Soc.*, **126**, 3016–3017.
19. Vainrub, A. and Pettitt, B.M. (2000) Thermodynamics of association to a molecule immobilized in an electric double layer. *Chem. Phys. Lett.*, **323**, 160–166.
20. Di Giusto, D.A. and King, G.C. (2004) Strong positional preference in the interaction of LNA oligonucleotides with DNA polymerase and proofreading exonuclease activities: implications for genotyping assays. *Nucleic Acids Res.*, **32**, e32.