Medium Scale Integration of Molecular Logic Gates in an Automaton

Joanne Macdonald,*,† Yang Li,†,§, Marko Sutovic,†,§ Harvey Lederman,†,§ Kiran Pendri,†,§ Wanhong Lu,†,§ Benjamin L. Andrews,‡ Darko Stefanovic,‡ and Milan N. Stojanovic†

National Chemical Bonding Center: Center for Molecular Cybernetics, Division of Experimental Therapeutics, Department of Medicine, Columbia University, New York, New York, and Department of Computer Science, University of New Mexico, Albuquerque, New Mexico

Received September 1, 2006; Revised Manuscript Received September 14, 2006

ABSTRACT

The assembly of molecular automata that perform increasingly complex tasks, such as game playing, presents an unbiased test of molecular computation. We now report a second-generation deoxyribozyme-based automaton, MAYA-II, which plays a complete game of tic-tac-toe according to a perfect strategy. In silicon terminology, MAYA-II represents the first "medium-scale integrated molecular circuit", integrating 128 deoxyribozyme-based logic gates, 32 input DNA molecules, and 8 two-channel fluorescent outputs across 8 wells.

Molecular computation and circuits engineering[1−27] using a "silicomimetic" approach is currently focused on building molecular networks analogous to electrical engineering designs. These networks consist of logic gates, which perform Boolean logical operations such as AND, NOT, and OR on one or more inputs to produce an output. While individual molecular gates and small networks have previously been constructed, these gates are yet to be integrated at higher levels of complexity. Such integration in electrical engineering arises from massive parallelism and interconnections, rather than fundamental component complexity. The ability to truly integrate molecular components remains crucial for the construction of next-generation molecular devices.[11,14]

The largest solution-phase molecular circuits previously considered include networks combining up to 20 logic modules.[11,20] On a similar scale, utilizing a full set of deoxyribozyme-based logic gates,[6,19,24] we have constructed solution-phase computing circuits such as a half-adder,[7] ligase-phosphodiesterase cascades,[19] and most recently a full-adder that comprises 7 logic gates in a single tube.[24] Systems of greater complexity include molecular automata,[4,8−10] which are capable of analyzing a series of human or environmental inputs in a meaningful fashion. An unbiased test of automaton construction is game playing, and we have focused on tic-tac-toe: one of the simplest games of perfect information,

and yet a surprisingly complex combinatorial problem, with 2.65×10^{103} nonlosing strategies for a complete version of tic-tac-toe.[28] To this end, we previously constructed a deoxyribozyme-based molecular automaton (MAYA, a *mo*lecular *a*rray of *Y*ES and *A*ND gates) that plays a simplified symmetry-pruned game of tic-tac-toe encompassing 19 permissible game plays, using an array of 23 logic gates distributed over 8 wells.[8]

We now report the development of the first solution-phase molecular assembly comprising over 100 molecular logic gates, which more than quadruples the complexity performed by any previous system. Expanding from our original automaton, MAYA-II is a second generation molecular automaton capable of playing a complete game of tic-tac-toe against a human opponent, and encompasses 76 permissible game plays. MAYA-II is more user-friendly than its predecessor, as it signals both players move in a two-color output system and imposes no constraints on the position of the human player's first move. However, similar to MAYA-I, MAYA-II is constructed from three classes of stem-loop controlled deoxyribozyme-based logic gates that are allosterically modulated by input oligonucleotides to produce fluorescent output signals (Figure 1):[6−8] (i) YESx gates are activated by a single input x; (ii) xANDy gates are activated in the presence of two inputs x and y; and (iii) xANDyAND-NOTz gates are activated in the presence of inputs x and y only if a third inhibiting input z is absent. To play MAYA-II, a set of deoxyribozyme-based logic gates are arranged according to a predetermined strategy in a 3×3 array within a 384-well assay plate (see Figure 2A). Oligonucleotide inputs encoding the human moves are added successively

* To whom correspondence should be addressed. E-mail: jm2236@columbia.edu. Address: 650 W 168th St, Box 84, New York, NY 10032. Tel: +1-212-342-5610. Fax: +1-212-305-3475.
† Columbia University.
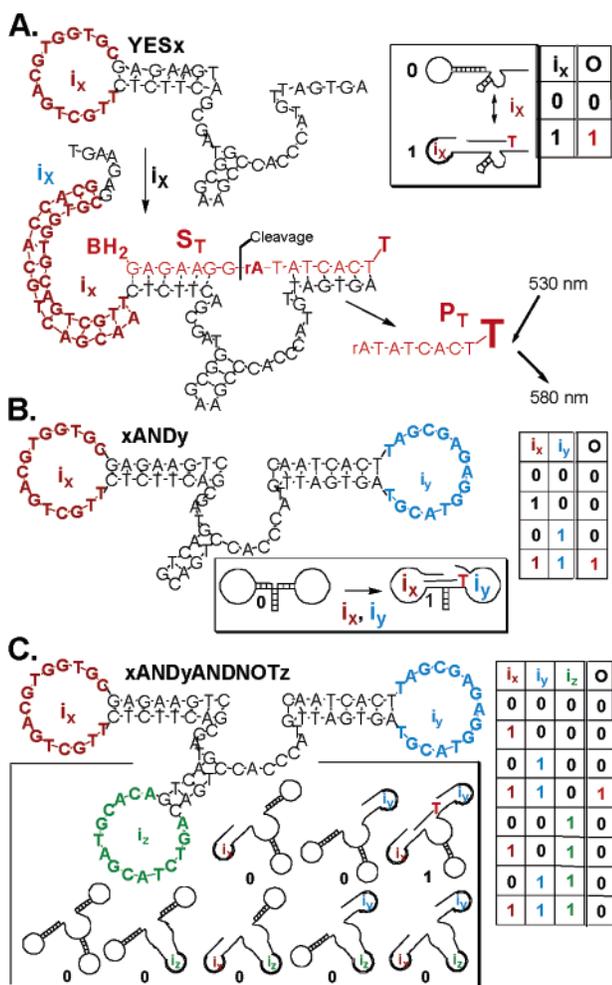‡ University of New Mexico.
§ NSF-funded high school internship program.

Figure 1. Automaton move gates and logic gate structures: (A) Automaton move gates were designed from E6 deoxyribozyme-based logic gates.[7,24,33] Upon addition of activating input, these gates cleave substrate S_T to produce P_T and an increase in TAMRA (T) fluorescence. YESx gates are activated by a single input x. (B) xANDy gates are activated in the presence of two inputs x and y. (C) xANDyANDNOTz gates are activated in the presence of inputs x and y only if a third inhibiting input "z" is not present. Inserts show a truth table of logic gate behavior.

to the wells and trigger the automaton's next move. After each input addition, the well-plate is analyzed using a fluorescent plate reader which follows accumulation of two fluorescent oligonucleotide outputs: display of the human move is observed in the "green" fluorescein output channel, and the automaton's response to the human input addition is observed in the "red" tetramethylrhodamine (TAMRA) output channel. An example game is shown in Figure 2B.

The game strategy for MAYA-II (Figure 2A and Supporting Information) is considerably different from MAYA-I. The automaton still moves first into the middle square (well 5) controlled by a constitutively active deoxyribozyme added immediately prior to the beginning of the game. However successive automaton moves are constructed as a hierarchical cascade of AND gates, with YES gates responding to the first human move (NOT loops are included to prevent secondary activation in already played wells or are redundant and included to minimize cumulative nondigital behavior in

side wells over several moves). In doing this, MAYA-II is a step toward programmable and generalizable MAYAs that are trainable to play any game strategy. Our strategy required 32 input oligonucleotides, encoding both the position and timing of a human move. These inputs are named I_{NM} wherein N is the position of the move (wells 1−4 or 6−9), and M is the timing of the move (1 for the first move, 2 for the second move, etc.). For example, input I_{62} would be added to every well when a human would like to indicate to the automaton they are moving into square 6 as their second move. This strategy was translated into Boolean logic amenable to deoxyribozyme-based implementation (Figure 2) using a custom computer program, resulting in 96 logic gates for automaton move calculations and an additional 32 logic gates to display human moves. We stress that, considering traditional difficulties in selecting nucleic acids suitable for computation[29,30] and the relatively high concentration of individual gates needed to accomplish readable outputs, it was far from certain at the outset of this project that such a number of inputs and gates could be coordinated in solution.

The 32 input oligonucleotide sequences (Table 1) were chosen to investigate both the inherent generality of our logic gate design, and our ability to derive inputs using computer assistance (akin to previous system designs[31]). In contrast to MAYA-I, where a smaller number of inputs meant trial-and-error substitution of inputs was feasible, we used an algorithm specifically devised for this purpose: (1) A theoretical library of stem-loop structures (containing a stem of 5 base-pairs and a loop of 15 nucleotides) was generated by applying a search algorithm, based on simple combinatorial constraints,[32] where loops containing stronger internal structures of more than two base-pairs were eliminated; (2) of the 10 795 generated sequences, a set of 32 15-mer loop sequences with no more than four nucleotides in common in a continuous stretch were selected for trial as oligonucleotide inputs and randomly assigned to human move and order positions; (3) these sequences were inserted into deoxyribozyme gate structures and analyzed using mfold;[32] (4) input sequences inducing gate misfolding were discarded and replaced with the next inputs from our library; (5) canonical gates and their reverse complement 15-mer input sequences were custom-synthesized and tested in solution-phase for digital gate behavior; (6) inputs and gates failing to show expected digital behavior were substituted with the next input from our collection. Tested input sequences are listed in Table 1. Out of the initial set of 32 inputs only three were rejected and substituted. Thus, while there is still space for improvement in the design of our algorithm, it led to minimization of trial-and-error from the input selection.

All automaton response gates were constructed from deoxyribozyme E6[7,24,33] (Figure 1), which cleaves oligonucleotide substrate S_T to produce product P_T and an increase in "red channel" TAMRA (T) fluorescence. Variable signal intensity was detected in some gates, and signal optimization was achieved by manipulating the 5′ or 3′ ends of the gate molecule, reversing the input loop sequences, or removing redundant NOT loops (summarized in Table 1, sequences

Figure 2. Schematic representation of MAYA-II. One-, two-, and three-input deoxyribozyme-based logic gates are allosterically modulated by 32 human-operated input oligonucleotides: 96 logic gates and one constitutively active deoxyribozyme distributed across nine wells calculate automaton moves, and 32 gates (boxed) display human moves by implementation of a two-color fluorogenic output system. Green loops denote positive regulation, and red loops denote negative regulation. (B) A representative game. Gates and active deoxyribozyme (final conditions listed in the Supporting Information) were mixed with 0.5 μM S_F, 1 μM S_T and dispensed into nine wells of a 384-well plate. Inputs (1 μM) were added in sequence into each well to signal the human players move, and both fluorescein (F) and TAMRA (T) fluorescence was measured every 15 min for 60 min in between each move. Results are expressed as the slope of signal increase over time (dF min^{-1}).

and graphs in Supporting Information). For human move visualization we used 8−17-based deoxyribozyme logic gates,[7,24,34] which cleave substrate S_F to produce product P_F and an increase in "green channel" fluorescein (F) fluorescence. However, only two input sequences were found to be active, and the underlying enzyme was re-engineered by

lengthening the substrate binding region (Figure 3A) leading to the perfect digital behavior of all inputs (Figure 3B). After individual testing and optimization, gates were mixed according to the MAYA-II algorithm (Figure 2) and retested for digital behavior to exclude the possibility of undesirable cross reactivity. Results were expressed as the slope of

Table 1. Input Sequences for MAYA-II and Modifications Required for Human Move Gates

inputs and sequence[a]	YES/M	5′ AND total	5′ AND Modified +	5′ AND Modified −	5′ AND Modified R	3′ AND total	3′ AND Modified +	3′ AND Modified −	3′ AND Modified R	NOT/x
I11 AACGACTGCACCACG	1	6								3/x3
I12 CTCTCCCTGTACCCA		2			2	5				
I13 ACCCCTCTCGCTCTT		1				6		−4	1	
I14 TTCTGCCTTGATCCG						1				
I21 TGTTGTCTTATCCAT	1	5		−4		1			1	3
I22 TCAGATGCTACGTGT		6				7			2	10
*I*23 ACCGTACTCGACCTA						3		−1		
I24 TCGGATCTCGGTTTC						1		−1		
I31 TACACGCTGGTCAAT	1	5		−4		1			1	3
I32 CACTATCTCGAATCA						7	+5			
I33 GCGTGACTGCGGCAT		1				6		−1	1	
I34 GTTGGTCTTGTAGGA						1				
I41 GCTAGGCTATCGCGT	1/-1	6								3
I42 TAATACCTGAGCGGG		7		−2		6	+2			10
I43 TACCCCCTAGTCTGC		1		−1	1	2		−1		
I44 AACGGACTTCAACAG		1		−1	1					
I61 CGGGATCTCGTCGGT	1	6	+2							3
I62 ATCGCTCTCCATGCA		8		−4		5				10/x10
I63 ATCTATCTCGTTCCG		2			2	1		−1		
I64 ACTCCGCTCGACTTA						1				
I71 GGATCACTTACGTAT	1/+1	6	+1							3
I72 GGTAGCCTTTTATCG						7	+4			
I73 CATTGCCTCGATATC		2				5		−1		
I74 CCAGACCTTTCAAGT						1				
I81 TGCGTACTTTGGGTC	1	6								3
I82 TCAGGGCTACGCAAG		7		−6		6	+1	−3	1	10
I83 TAATTACTGTTTCAC		1				2	+2			
I84 GGATGCCTGGCGTCT		1		−1	1					
I91 TGCTATCTCGACAAG	1	6	+1							3
I92 CTCAGGCTGTGTATT						7	+4			
I93 CAGAGCTATACGGAG		2				5		−1		
I94 GCTACTCTGGGTGCT						1				
Totals:	8/+1/−1	88	+4	−23	7	88	+18	−14	7	64/x13

fluorescence change over time (ΔF min^{-1}, Figure 2). Automaton move gates with slopes greater than 3000 ΔF min^{-1} were considered positive, and signals smaller than 2000 δF min^{-1} were discounted as background noise. Human move gates tended to give higher signals and rates of reaction over time and were thus included at lower concentrations. Even so, these gates typically gave a positive reaction of greater than 5000 ΔF min^{-1}, and reactions smaller than 4000 ΔF min^{-1} were considered noise. While most gates were not adversely affected within the mixtures, variable signal intensity was observed for some automaton move gates, and these were optimized either as described above, or by altering the concentration of individual gates within the mix. Interestingly, a directly proportional signal to concentration ratio was not observed for every gate, as some gates were inhibited by increasing concentrations within the mix (Supporting Information).

Upon establishment of the final conditions (Supporting Information), MAYA-II was constructed as a set of eight

tubes (the Well 5 tube containing active deoxyribozyme was sometimes omitted), and all 76 tic-tac-toe games were repeatedly tested. MAYA-II was able to play perfectly a general tic-tac-toe game by successfully signaling both human and automaton moves. Small immediate increases of fluorescence upon input addition, most likely the result of a conformational change of a gate complexed with substrate, were occasionally observed at the first measurement (the first 15 min of reaction); however this was distinguishable from positive signals as the fluorescence did not continue to increase. Thus, digital behavior could be reliably confirmed within 30 min of input addition. An example game is shown in Figure 2, and the results from all 76 games are provided in the Supporting Information.

The success of MAYA-II indicates the maturity of our deoxyribozyme-based logic gate system as a "plug and play" integrated logic gate system. MAYA-II integrates 128 molecular logic gates, 32 oligonucleotide inputs, and 8 two-channel fluorescent outputs across 8 wells. It could be argued
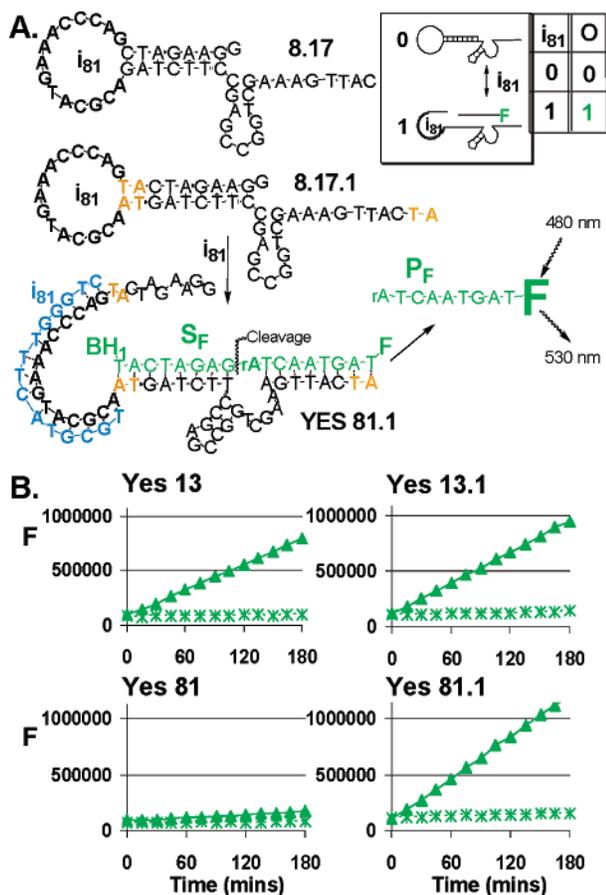
Figure 3. Human move gates: (A) Human move gates were designed from 8.17 deoxyribozyme-based logic gates.[7,24,34] Upon addition of activating input, these gates cleave substrate S_F to produce product P_F and an increase in fluorescein fluorescence (F). Lengthening of the substrate binding region created 8.17.1, with more reliable enzyme activity. The insert shows a truth table of logic gate behavior. (B) Raw fluorescence activity (F) of 8.17 and 8.17.1-based gates (20 nM) using 1 μM S_F in the presence (triangles) or absence (stars) of 1 μM inputs. Gates derived from 8.17 were variably active, as demonstrated here by Yes 13 (active) and Yes 81 (not active), whereas gates derived from 8.17.1 were always active, as displayed for Yes 13.1 and Yes 81.1

that by integrating more than 100 molecular logic gates in a single system, MAYA-II represents the first "medium-scale integrated molecular circuit" in solution. This increased complexity of MAYA-II has enabled refinement of our deoxyribozyme logic gate model, allowing the development of design principles for optimizing digital gate behavior[35] and the generation of a library of known input sequences (Table 1). Because our gates are made from DNA, we also expect them to be amenable to evolutionary methods for development.[36] Our symmetrical game strategy enabled the entire game to be essentially encoded as a series of YES and AND gates, which take into account only two human moves: the current and preceding. We propose that a total of 152 gates could be used to encode any symmetrical game strategy into any automaton using the above-defined 32 inputs and allowing for subsequent additional activation in already played wells. We are currently constructing the full set of gates for the development of selection procedures to obtain fully trainable automata.

The practical applications of this massive parallel integration are more likely in oligonucleotide analysis rather than competition with silicon in high-speed computing. For example, the ability to detect and analyze combinations of multiple DNA sequences within minutes has direct applications in microarray style diagnostics. Automata the size of MAYA-II analyze the space of 2^{32} possible subsets of the 32 input oligonucleotides and partition it into equivalence classes signaled by unique two-color, eight-well patterns, for a total of up to $2^{16} = 65\,536$ patterns. Based on MAYA-II, we are currently developing several systems for multiplex SNP detection and viral lineage attribution. Moreover, the versatility of the input and output system allows coupling of logic gate processing to both upstream and downstream events, such as the detection and release of small molecules and the inhibition of enzymatic activity.[37] We are investigating the depth to which serial connectivity can be achieved and are considering a reset function to allow gates to perform multiple tasks. These developments should allow for the application of deoxyribozyme logic gate technology in bidirectional signaling events and pave the way for the next generation of fully autonomous molecular devices.

Supporting Information Available: Detailed experimental methods, additional figures, all logic gate sequences, and results from all 76 games. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Adleman, L. M. *Science* **1994**, *266*, 1021−4.
(2) Elowitz, M. B.; Leibler, S. *Nature* **2000**, *403*, 335−8.
(3) Mao, C.; LaBean, T. H.; Relf, J. H.; Seeman, N. C. *Nature* **2000**, *407*, 493−6.
(4) Benenson, Y.; Paz-Elizur, T.; Adar, R.; Keinan, E.; Livneh, Z.; Shapiro, E. *Nature* **2001**, *414*, 430−4.
(5) Braich, R. S.; Chelyapov, N.; Johnson, C.; Rothemund, P. W.; Adleman, L. *Science* **2002**, *296*, 499−502.
(6) Stojanovic, M. N.; Mitchell, T. E.; Stefanovic, D. *J. Am. Chem. Soc.* **2002**, *124*, 3555−61.
(7) Stojanovic, M. N.; Stefanovic, D. *J. Am. Chem. Soc.* **2003**, *125*, 6673−6.
(8) Stojanovic, M. N.; Stefanovic, D. *Nat. Biotechnol.* **2003**, *21*, 1069−74.
(9) Adar, R.; Benenson, Y.; Linshiz, G.; Rosner, A.; Tishby, N.; Shapiro, E. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9960−5.
(10) Benenson, Y.; Gil, B.; Ben-Dor, U.; Adar, R.; Shapiro, E. *Nature* **2004**, *429*, 423−9.
(11) de Silva, A. P.; McClenaghan, N. D. *Chemistry* **2004**, *10*, 574−86.
(12) Rothemund, P. W.; Papadakis, N.; Winfree, E. *PLoS Biol.* **2004**, *2*, e424.
(13) Schmidt, K. A.; Henkel, C. V.; Rozenberg, G.; Spaink, H. P. *Nucleic Acids Res.* **2004**, *32*, 4962−8.
(14) Simpson, M. L. *Trends Biotechnol.* **2004**, *22*, 555−7.
(15) Su, X.; Smith, L. M. *Nucleic Acids Res.* **2004**, *32*, 3115−23.
(16) Szacilowski, K. *Chemistry* **2004**, *10*, 2520−8.
(17) Benner, S. A.; Sismour, A. M. *Nat. Rev. Genet.* **2005**, *6*, 533−43.

(18) Grover, W. H.; Mathies, R. A. *Lab. Chip* **2005**, *5*, 1033−40.

(19) Stojanovic, M. N.; Semova, S.; Kolpashchikov, D.; Macdonald, J.; Morgan, C.; Stefanovic, D. *J. Am. Chem. Soc.* **2005**, *127*, 6914−5.

(20) de Silva, A. P.; Leydet, Y.; Lincheneau, C.; McClenaghan, N. D. *J. Phys.: Condens. Matter* **2006**, *18*, S1847−72.

(21) Henkel, C. V.; Back, T.; Kok, J. N.; Rozenberg, G.; Spaink, H. P. *Biosystems* **2006**.

(22) Isaacs, F. J.; Dwyer, D. J.; Collins, J. J. *Nat. Biotechnol.* **2006**, *24*, 545−54.

(23) Komiya, K.; Sakamoto, K.; Kameda, A.; Yamamoto, M.; Ohuchi, A.; Kiga, D.; Yokoyama, S.; Hagiya, M. *Biosystems* **2006**, *83*, 18−25.

(24) Lederman, H.; Macdonald, J.; Stefanovic, D.; Stojanovic, M. N. *Biochemistry* **2006**, *45*, 1194−99.

(25) Liu, Y.; Jiang, W.; Zhang, H. Y.; Li, C. J. *J. Phys. Chem. B* **2006**, *110*, 14231−35.

(26) Magri, D. C.; Brown, G. J.; McClean, G. D.; de Silva, A. P. *J. Am. Chem. Soc.* **2006**, *128*, 4950−1.

(27) Penchovsky, R.; Breaker, R. R. *Nat. Biotechnol.* **2005**, *23*, 1424−33.

(28) Andrews, B., Games, Strategies, and Boolean Formula Manipulation. In *Dept. of Computer Science, Univ. of New Mexico, UNM CS Tech. Rep. TR-CS-2005*−41; University of New Mexico: NM, 2005.

(29) Mauri, G.; Ferretti, C. In *Word design for molecular computing: A survey*, 9th International Workshop on DNA-Based Computers, 2003; Springer: 2004; pp 37−46.

(30) Sager, J.; Stefanovic, D. In *Designing Nucleotide Sequences for Computation: A Survey of Constraints,* 11th International Workshop on DNA-Based Computers, 2005; Springer: 2006; pp 275−289.

(31) Seeman, N. C. *J. Theor. Biol.* **1982**, *99*, 237−47.

(32) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406−15.

(33) Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1995**, *2*, 655−60.

(34) Santoro, S. W.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4262−6.

(35) Macdonald, J.; Stefanovic, D.; Stojanovic, M. N. Solution-Phase Molecular-Scale Computation With Deoxyribozyme-Based Logic Gates and Fluorescent Readouts. In *Fluorescent energy transfer nucleic acid probes: designs and protocols*; Didenko, V. V., Ed.; Humana Press Inc.: Totowa, NJ, 2006; Vol. 335, pp 343−363.

(36) Wright, M. C.; Joyce, G. F. *Science* **1997**, *276*, 614−617.

(37) Kolpashchikov, D. M.; Stojanovic, M. N. *J. Am. Chem. Soc.* **2005**, *127*, 11348−51.