DNA Computation

DOI: 10.1002/anie.201406892

Interfacing Synthetic DNA Logic Operations with Protein Outputs**

*Alexander Prokup and Alexander Deiters**

Abstract: *DNA logic gates are devices composed entirely of DNA that perform Boolean logic operations on one or more oligonucleotide inputs. Typical outputs of DNA logic gates are oligonucleotides or fluorescent signals. Direct activation of protein function has not been engineered as an output of a DNA-based computational circuit. Explicit control of protein activation enables the immediate triggering of enzyme function and could yield DNA computation outputs that are otherwise difficult to generate. By using zinc-finger proteins, AND, OR, and NOR logic gates were created that respond to short oligonucleotide inputs and lead to the activation or deactivation of a split-luciferase enzyme. The gate designs are simple and modular, thus enabling integration with larger multigate circuits, and the modular structure gives flexibility in the choice of protein output. The gates were also modified with translator circuits to provide protein activation in response to microRNA inputs as potential cellular cancer markers.*

Similar to their silicon-based counterparts, DNA logic gates are fundamental computational devices that generate a specific output based on oligonucleotide inputs. Logic gates are being used for increasingly complex applications, such as simulating human memory,[1] playing games,[2] computing mathematical addition,[3–5] and detecting microRNA patterns.[6] A limitation in the current DNA-based gate designs is a lack of direct interfacing with biological components other than oligonucleotides. We report the first direct integration of DNA logic gates with protein outputs. By linking DNA computation to protein activity, a wide range of biological events could be triggered based purely on oligonucelotide inputs.

Our approach makes use of zinc-finger proteins since they are able to bind DNA without disrupting the structure of the logic gate and can be easily fused to split-protein components. Zinc-finger proteins are naturally occurring, sequence-specific DNA binders that have been modified with various effector domains[7], for example, to regulate transcription,[8] manipulate mitochondrial DNA,[9] or inhibit viral replication.[10] A single zinc finger will recognize an arrangement of three nucleotides. Generally, six fingers are fused together to create a larger protein that can identify a unique 18-nucleotide sequence. The most common group of zinc fingers is based on the Cys_2His_2 motif[11] and includes the proteins AaRT and E2C. The majority of zinc-finger proteins, such as E2C,[12] recognize guanine-rich DNA sequences, however, the protein AaRT was specifically designed to bind adenine-rich sequences.[13] Changing a single nucleotide in the binding site has been shown to negatively affect the binding of zinc fingers to DNA by increasing the K_D value more than 100-fold.[14] The utilization of two unique zinc-finger proteins, AaRT and E2C, in our logic-gate design ensures sequence-specific protein activation.

In addition to the zinc-finger component, the developed logic gates contain a split-luciferase enzyme to generate a luminescence readout, which gives a highly sensitive biosensor design.[15] However, the function of a wide range of other split-proteins, including green fluorescent protein,[16] β-galactosidase,[17] and TEV protease,[18] could be triggered as well. A split-protein consists of two halves of the original protein. Separately, each half is inactive and does not exhibit enzymatic activity. When the two complementary halves are brought into close proximity, the enzyme structure is reconstituted and activity is restored. In our DNA logic-gate design, the N-terminal half of luciferase is fused to AaRT while the C-terminal half is fused to E2C. Although each gate contains the zinc finger and split-luciferase fusion protein, the operation of each gate is inherently unique.

In order to demonstrate the triggering of protein function by the DNA logic gate, we selected an AND gate (Figure 1 A) as an initial example. This gate only produces an output when both inputs are present. The zinc finger AND gate follows the most simple design of our three gates: it consists of a single strand of DNA (G_{AND}) and is activated by two inputs (A and B). As shown in Figure 1 B, the two zinc-finger binding sites are only completed after the hybridization of both inputs to the gate strand. After binding of the zinc-finger proteins to the DNA scaffold, the split-luciferase halves are able to establish a functional luciferase enzyme. Therefore, activation of the AND gate only occurs in the presence of both inputs, as observed when A and B were added to the G_{AND} strand (Figure 1 C). Only minimal baseline luminescence activity was detected in the presence of just A or just B; however, a 5-fold increase in luciferase activity was observed when both inputs were present. Owing to the stringent requirements for an AND gate, it is often used for the "carry" function in larger devices, like the half-adder.[19]

In contrast to the AND gate, an OR gate (Figure 2 A) yields an output in the presence of either of the two inputs. The corresponding DNA-based circuitry is generated from two preformed duplexes (Figure 2 B): one consisting of an incomplete AaRT binding site ($G_{OR,A}$) and the other containing the incomplete E2C binding site ($G_{OR,B}$). When input A or B is added and hybridized to $G_{OR,A}$ or $G_{OR,B}$, the missing

[*] A. Prokup, Prof. Dr. A. Deiters
University of Pittsburgh, Department of Chemistry
Pittsburgh, PA 15260 (USA)
E-mail: deiters@pitt.edu

[**] This work was supported by the University of Pittsburgh. We acknowledge Dr. Chungjung Chou for assistance in developing protein expression conditions and Dr. Indraneel Ghosh (University of Arizona) for the expression constructs.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201406892.

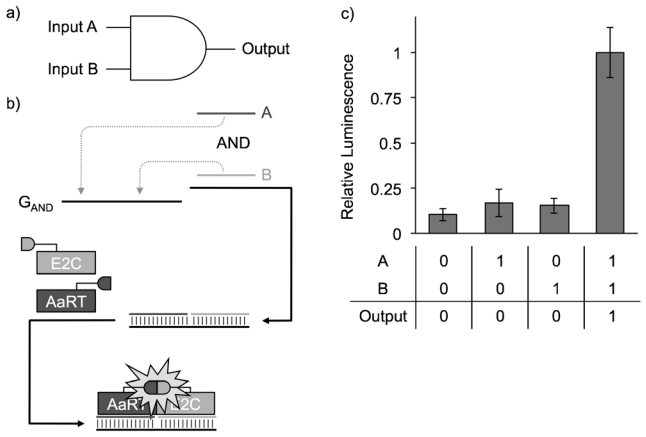

Figure 1. a) The electrical symbol for the AND gate. b) A scheme of the AND gate based on DNA–protein interactions. Split-luciferase halves are represented by solid half ovals, the zinc-finger proteins by rectangles, and DNA is shown as solid lines. c) Luminescence readouts are shown for the AND gate after the inputs A and B were added in different combinations to the G_{AND} strand. The AND gate truth table is displayed below the chart. Three independent experiments were averaged and the error bars represent the standard deviation.

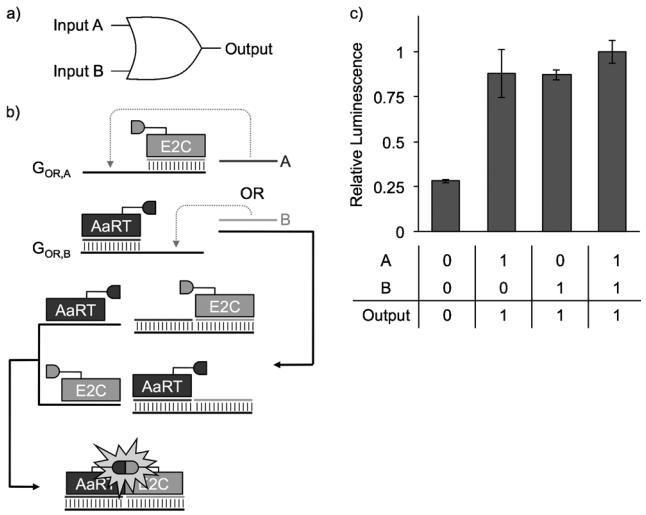

Figure 2. a) The electrical symbol for the OR gate. b) A scheme of the OR gate, based on DNA–protein interactions. Split-luciferase halves are represented by half ovals, the zinc-finger proteins by rectangles, and DNA is shown as solid lines. c) Luminescence readouts are shown for the OR gate after the two inputs A and B were added in different combinations to the G_{OR} duplexes. The OR gate truth table is displayed below the chart. Three independent experiments were averaged and the error bars represent the standard deviation.

DNA binding site for the zinc-finger is formed. The second zinc finger is then able to bind to the DNA, thereby allowing the formation of an active luciferase enzyme. Thus, only a single input is necessary to activate this DNA-based OR gate (Figure 2C). As expected, when either A or B was added, a luminescence signal was produced. In the absence of any input strand, only basal levels of luciferase activity were detected. Slightly higher background activity was observed for the OR gate in comparison to the other gates. This effect may result from the doubled zinc-finger concentration, which is required owing to the use of two preformed duplexes, $G_{OR,A}$ and $G_{OR,B}$. However, a clear and significant difference in luminescence was detected in the presence of the inputs. OR gates are commonly found in electrical devices, including simple multiplexers, which can function as electronic rotary switches.[20]

As a third, essential logic gate operation, the NOR gate (Figure 3A) functions as an inverse (or negated) OR gate.

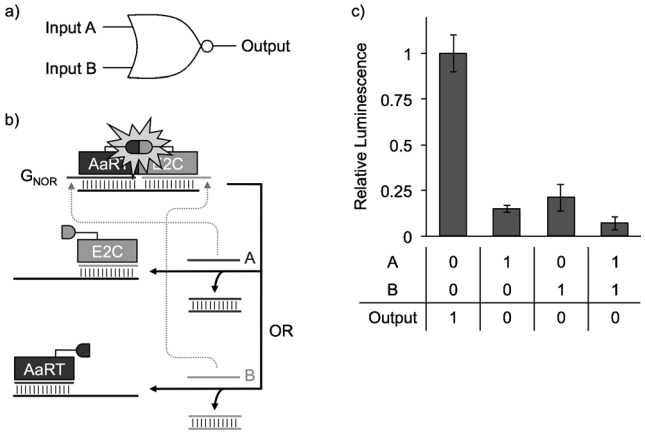

Figure 3. a) The electrical symbol for the NOR gate. b) A scheme of the NOR gate based on DNA–protein interactions. Split-luciferase halves are represented by half ovals, the zinc-finger proteins by rectangles, and DNA is shown as solid lines. c) Luminescence readouts are shown for the NOR gate after the two inputs A and B were added in different combinations to the G_{NOR} duplex. The NOR gate truth table is displayed below the chart. Three independent experiments were averaged and the error bars represent the standard deviation.

Thus, the presence of either input signal will lead to no output signal. Translated into DNA circuitry interfaced with protein outputs, this occurs when either input A or B binds to a toehold region of the NOR gate, thereby displacing the A or B strand (Figure 3B). With the two zinc fingers far apart, the split-luciferase halves are separated and do not produce a luminescence output. A minimum five-fold dynamic range can be observed with the addition of either input, thus distinguishing the high and low Boolean outputs (Figure 3C). A NOR gate is especially useful since it is often considered a universal gate; any other logic gate can be created through a combination of multiple NOR gates.[21] For example, a XOR gate can be created from five NOR gates through serial and parallel connections.

Connecting multiple logic gates in a larger network[22] enables the construction of more complex devices, like half-adders[19] and multiplexers.[20] The final Boolean computation in any of these networks could be readily performed by the zinc-finger-interfaced gates presented here. We constructed a sub-network by connecting an AND gate and a NOR gate in series (Figure 4A). When both inputs B and C are present, the AND gate will release an output, which acts as an input for the NOR gate. Either the AND gate output or the input A will trigger the NOR gate and eliminate a zinc-finger binding

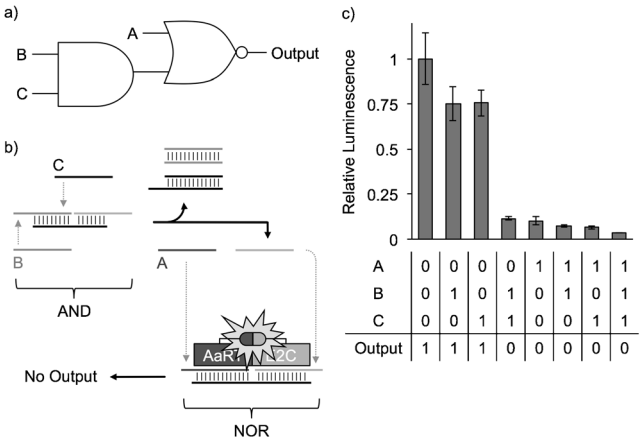

Figure 4. a) An AND gate and a NOR gate connected in series. b) A scheme of the corresponding DNA computation circuit. The ssDNA output from the AND gate becomes an input for the NOR gate. The final output from the NOR gate is emitted as luminescence. c) Luminescence data for the AND–NOR circuit. Three independent experiments were averaged and the error bars represent the standard deviation.

site (Figure 4B). Luminescence measurements confirmed the expected activity of the network (Figure 4C). Consistent with the truth table for the circuit, high luminescence values are only observed in the absence of any input or in the presence of only B or only C.

The encoding of multiple logic gates by simple DNA frameworks combined with zinc-finger proteins demonstrates the versatility of protein-interfaced DNA components. Obviously, a constraint is the strict sequence specificity required for zinc-finger proteins to bind. However, this limitation can be overcome by simply employing translator gates to convert any oligonucleotide sequence into the sequence of DNA inputs like A or B.[6a] As a proof of principle, we designed translator gates for the microRNAs miR-21 and miR-122, which are implicated in several diseases.[23–25] NOR gate input toe-holds are sequestered inside DNA duplexes, and in the absence of miRNAs, no decrease in luminescence output is observed (Figure S1 in the Supporting Information). The addition of the miRNAs to the translator gates releases a NOR gate input with an exposed toe-hold binding site. The inputs are then able to interact with the NOR gate as shown in Figure 4, thereby leading to a reduction in the luminescence output. The OR, AND, and NOR logic gates presented herein thus represent modular biological devices that can be used in the final layer of DNA computation circuits[1,26] in order to interface them with protein outputs.

DNA logic gates typically yield single-stranded DNA output.[27–29] Luminescence outputs have not been incorporated into logic gate designs before and may show advantages for reducing background signal,[30] as well as for expanding the range of applications for DNA computational devices. The activation of an enzyme in response to a DNA computation event has the potential to provide a facile solution to the often absent signal amplification in DNA devices.[31] Moreover, the split-luciferase could be replaced by any number of split-proteins, which greatly expands the output possibilities to more than luminescence, including direct activation of a protein of interest and the triggering of a wide range of biological processes by either activation (AND and OR gates) or deactivation (NOR gate) of a split-protein, such as a protease (TEV),[18] recombinase (Cre),[32] hydrolase (β-galactosidase),[17] or fluorophore (GFP).[16] A simple DNA circuit output could thus lead to significant modification and perturbation of the biological system it is employed in.

Experimental Section

NOR gate operation: White-walled, clear-bottomed 96-well plates and TE/Mg^{2+} buffer (10 mM tris-HCl pH 8, 100 mM ethylenediaminetetraacetic acid, 12.5 mM $MgCl_2$) were used. The total volume of the logic gate experiments was 50 μL before the addition of 50 μL luminescence assay reagents. The G_{NOR} duplex (200 nM) was added to all possible combinations of inputs A and B (400 nM each) to wells containing AaRT and MBP-E2C (200 nM). To assay for luminescence, the Bright-Glo Luciferase Assay System (50 μL) was added after incubation (room temperature, 45 min) with all other components. Luminescence was detected by using a BioTek Synergy 4 plate reader. Triplicate experiments were performed for each condition.

Received: July 4, 2014
Published online: October 3, 2014

Keywords: DNA structures · DNA recognition · logic gates · oligonucleotides · proteins

[1] L. Qian, E. Winfree, J. Bruck, *Nature* **2011**, *475*, 368–372.
[2] M. N. Stojanovic, D. Stefanovic, *Nat. Biotechnol.* **2003**, *21*, 1069–1074.
[3] R. Orbach, F. Wang, O. Lioubashevski, R. D. Levine, F. Remacle, I. Willner, *Chem. Sci.* **2014**, *5*, 3381–3387.
[4] J. Elbaz, O. Lioubashevski, F. Wang, F. Remacle, R. D. Levine, I. Willner, *Nat. Nanotechnol.* **2010**, *5*, 417–422.
[5] H. Lederman, J. Macdonald, D. Stefanovic, M. N. Stojanovic, *Biochemistry* **2006**, *45*, 1194–1199.
[6] a) G. Seelig, D. Soloveichik, D. Y. Zhang, E. Winfree, *Science* **2006**, *314*, 1585–1588; b) J. Hemphill, A. Deiters, *J. Am. Chem. Soc.* **2013**, *135*, 10512–10518.
[7] R. Gabriel, A. Lombardo, A. Arens, J. C. Miller, P. Genovese, C. Kaeppel, A. Nowrouzi, C. C. Bartholomae, J. Wang, G. Friedman, M. C. Holmes, P. D. Gregory, H. Glimm, M. Schmidt, L. Naldini, C. von Kalle, *Nat. Biotechnol.* **2011**, *29*, 816–823.
[8] Y. Choo, I. Sánchez-García, A. Klug, *Nature* **1994**, *372*, 642–645.
[9] M. Minczuk, M. A. Papworth, P. Kolasinska, M. P. Murphy, A. Klug, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19689–19694.
[10] L. Reynolds, C. Ullman, M. Moore, M. Isalan, M. J. West, P. Clapham, A. Klug, Y. Choo, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1615–1620.
[11] J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **1985**, *4*, 1609–1614.
[12] R. R. Beerli, U. Schopfer, B. Dreier, C. F. Barbas, *J. Biol. Chem.* **2000**, *275*, 32617–32627.
[13] D. J. Segal, J. W. Crotty, M. S. Bhakta, C. F. Barbas, N. C. Horton, *J. Mol. Biol.* **2006**, *363*, 405–421.
[14] E. J. Rebar, C. O. Pabo, *Science* **1994**, *263*, 671–673.
[15] a) B. W. Jester, K. J. Cox, A. Gaj, C. D. Shomin, J. R. Porter, I. Ghosh, *J. Am. Chem. Soc.* **2010**, *132*, 11727–11735; b) T. Hashimoto, K. W. Adams, Z. Fan, P. J. McLean, B. T. Hyman, *J. Biol. Chem.* **2011**, *286*, 27081–27091; c) J. L. Furman, A. H. Badran, O. Ajulo, J. R. Porter, C. I. Stains, D. J. Segal, I. Ghosh, *J. Am. Chem. Soc.* **2010**, *132*, 11692–11701; d) I. Ghosh, C. I.

Stains, A. T. Ooi, D. J. Segal, *Mol. BioSyst.* **2006**, *2*, 551 – 560; e) S. S. Shekhawat, J. R. Porter, A. Sriprasad, I. Ghosh, *J. Am. Chem. Soc.* **2009**, *131*, 15284 – 15290.
[16] C. I. Stains, J. R. Porter, A. T. Ooi, D. J. Segal, I. Ghosh, *J. Am. Chem. Soc.* **2005**, *127*, 10782 – 10783.
[17] A. M. Broome, N. Bhavsar, G. Ramamurthy, G. Newton, J. P. Basilion, *Mol. Pharmceutics* **2010**, *7*, 60 – 74.
[18] M. C. Wehr, R. Laage, U. Bolz, T. M. Fischer, S. Grünewald, S. Scheek, A. Bach, K. A. Nave, M. J. Rossner, *Nat. Methods* **2006**, *3*, 985 – 993.
[19] M. N. Stojanović, D. Stefanović, *J. Am. Chem. Soc.* **2003**, *125*, 6673 – 6676.
[20] A. Credi, *Angew. Chem. Int. Ed.* **2007**, *46*, 5472 – 5475; *Angew. Chem.* **2007**, *119*, 5568 – 5572.
[21] J. Zhou, M. A. Arugula, J. Halámek, M. Pita, E. Katz, *J. Phys. Chem. B* **2009**, *113*, 16065 – 16070.
[22] L. Qian, E. Winfree, *Science* **2011**, *332*, 1196 – 1201.
[23] P. P. Medina, M. Nolde, F. J. Slack, *Nature* **2010**, *467*, 86 – 90.
[24] C. Coulouarn, V. M. Factor, J. B. Andersen, M. E. Durkin, S. S. Thorgeirsson, *Oncogene* **2009**, *28*, 3526 – 3536.
[25] C. Jopling, *RNA Biol.* **2012**, *9*, 137 – 142.
[26] J. Macdonald, Y. Li, M. Sutovic, H. Lederman, K. Pendri, W. Lu, B. L. Andrews, D. Stefanovic, M. N. Stojanovic, *Nano Lett.* **2006**, *6*, 2598 – 2603.
[27] W. Yoshida, Y. Yokobayashi, *Chem. Commun.* **2007**, 195 – 197.
[28] L. Gardner, A. Deiters, *Curr. Opin. Chem. Biol.* **2012**, *16*, 292 – 299.
[29] A. Prokup, J. Hemphill, A. Deiters, *J. Am. Chem. Soc.* **2012**, *134*, 3810 – 3815.
[30] A. Roda, P. Pasini, M. Mirasoli, E. Michelini, M. Guardigli, *Trends Biotechnol.* **2004**, *22*, 295 – 303.
[31] a) X. Chen, N. Briggs, J. R. McLain, A. D. Ellington, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5386 – 5391; b) Y. Zhang, Z. Tang, J. Wang, H. Wu, A. Maham, Y. Lin, *Anal. Chem.* **2010**, *82*, 6440 – 6446.
[32] N. Jullien, F. Sampieri, A. Enjalbert, J. P. Herman, *Nucleic Acids Res.* **2003**, *31*, e131.

