Programming DNA-Based Biomolecular Reaction Networks on Cancer Cell Membranes

Journal:	*Journal of the American Chemical Society*
Manuscript ID	ja-2019-05598u.R2
Manuscript Type:	Communication
Date Submitted by the Author:	05-Oct-2019
Complete List of Authors:	Song, Tianqi; Duke University, Computer Science Shah, Shalin; Duke University, Electrical & Computer Engineering Bui, Hieu; Duke University, Computer Science Garg, Sudhanshu; Duke University, COMPUTER SCIENCE Eshra, Abeer; Duke University, Computer Science Fu, Daniel; Duke University Yang, Ming; Duke University Mokhtar, Reem; Duke University, Computer Science Reif, John; Duke University, Dept. of Computer Science

Programming DNA-Based Biomolecular Reaction Networks on Cancer Cell Membranes

Tianqi Song,[†] Shalin Shah,[‡] Hieu Bui,[¶] Sudhanshu Garg,[†] Abeer Eshra,[†,§] Daniel Fu,[†] Ming Yang,[†] Reem Mokhtar,[†] and John Reif*[,†,‡]

[†]*Department of Computer Science, Duke University, Durham, North Carolina 27708, United States*
[‡]*Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708, United States*
[¶]*National Research Council, 500 Fifth Street NW, Keck 576, Washington, DC 20001, United States*
[§]*Department of Computer Science and Engineering, Faculty of Electronic Eng., Menoufia University, Menouf, Menoufia 32831, Egypt*

Received October 5, 2019; E-mail: reif@cs.duke.edu

Abstract: DNA is a highly programmable biomolecule and has been used to construct biological circuits for different purposes. An important development of DNA circuits is to process the information of receptors on cell membranes. In this paper, we introduce an architecture to program localized DNA-based biomolecular reaction networks on cancer cell membranes. Based on our architecture, various types of reaction networks have been experimentally demonstrated, from simple linear cascades to reaction networks of complex structures. These localized DNA-based reaction networks can be used for medical applications such as cancer cell detection. Compared to prior work of DNA circuits for evaluating cell membrane receptors, the DNA circuits by our architecture have several major advantages including simpler design, lower leak, lower cost, and higher signal-to-background ratio.

DNA is a very promising biomolecule for constructing biological circuits for different aims.[1–30] Particularly, DNA circuits can be used to evaluate the information of receptors on cancer cell membranes for medical applications such as cancer cell detection, and impressive work has been done on this topic.[4,6,8,9,27,30]

However, most prior work has at least one of the following drawbacks that should be improved: Firstly, the circuits are not fully localized,[9] which means that it cannot be guaranteed that the inputs to a circuit are from a single cell. This can cause false-positive result for applications like cancer detection in an environment with multiple cell types. Secondly, the basic reaction motif of the circuits is made from two or several DNA strands.[4,6,8,9,27] Compared with single-stranded motifs such as DNA hairpin (stem-loop), motifs by multiple DNA strands need more design and operation work, and also are more prone to leak caused by imperfection in stoichiometry, annealing and purification. Thirdly, a DNA gate is permanently conjugated to an aptamer or antibody that targets a designated cell membrane receptor,[4,6,8,9,27,30] which makes it difficult and costly to reprogram a DNA gate to target a different receptor. Therefore, there is a need to develop new architectures to conquer such drawbacks.

In this paper, we introduce an architecture to program DNA-based biomolecular reaction networks on cancer cell membranes to conquer the drawbacks mentioned before. Specifically, a reaction network by our architecture consists of multiple nodes that can interact with each other. The nodes of a network target designated cancer cell membrane receptors using DNA aptamers, and are fully localized on the membrane. If all targeted receptors exist on a cancer cell membrane, the circuit will be complete and able to work (Figure 1). The basic reaction motif is based on DNA hairpin which is a single-stranded structure, and then needs much less design and operation work compared to multi-strand structures. We have demonstrated both linear cascade reactions and more complex reaction networks that can perform sophisticated logic computation. A flexible addressing mechanism is used such that a node in our architecture can be easily reprogrammed to target different cancer cell membrane receptors. We have also demonstrated a prototype of cancer cell detection using a reaction network by our architecture.

The nodes in a reaction network have a modular design as shown in Figure 1 (a). Each node has a reaction module and an addressing module. The reaction module is responsible for cascade reactions, and the addressing module is responsible for targeting a receptor using a DNA aptamer. A DNA hairpin is used as the motif for the reaction module for several reasons: Firstly, it is a single-stranded DNA structure, which is much easier to design and engineer compared to structures by multiple DNA strands. Secondly, it is more resilient to leak caused by imperfect stoichiometry, purification and annealing compared to structures by multiple DNA strands. Thirdly, it is easy to cascade into complex reaction networks.[1] The reaction module and addressing module are connected by DNA hybridization between two complementary domains. A node is made by simply mixing the two modules by 1:1 concentration ratio. When it needs to reuse a reaction module to target a different receptor, we just take a new addressing module that has the corresponding aptamer, and mix it with the prior reaction module. This property of our architecture makes it easy and cheap to reprogram a node to target different receptors.

For explaining how to build a reaction network from nodes, an example of 2-layer linear cascade is shown in Figure 1 (b). There are two nodes in the network that target different cell membrane receptors using different aptamers. The initiator reacts with node A by DNA strand displacement, and the output of node A reacts with node B. The output of node B reacts with the reporter complex such that the fluorophore is tethered to the cell membrane. The nodes are mixed with cells first. The initiator and reporter complex are added later after filtering out the free nodes in the solution. Therefore, the reaction can happen only if both receptors present on the cell membrane. For the convenience

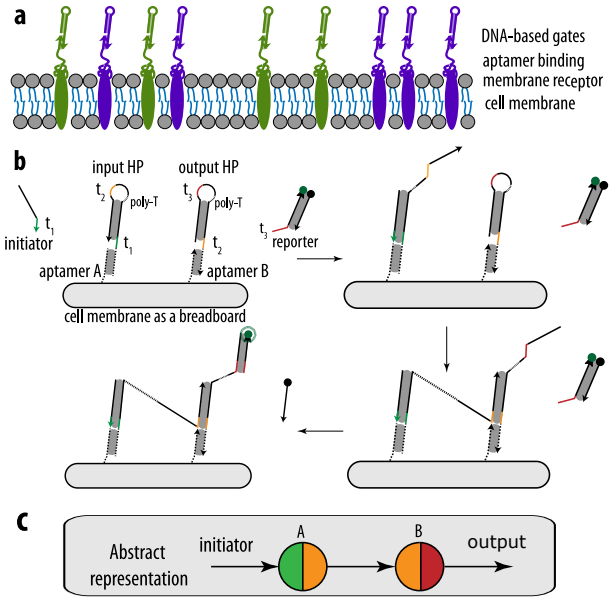

Figure 1. A brief description of our architecture. (a) A 2-layer linear cascade reaction network by our architecture. There are two types of nodes in the network that are indicated by two different colors. Each node has a reaction module (a DNA hairpin (HP)) and an addressing module (a single DNA strand containing a DNA aptamer), where a DNA aptamer (the curved part) is a DNA sequence that can be rationally designed[31] and recognize a particular cell membrane receptor which can range from small molecules to proteins.[32–35] The reaction module and addressing module are connected by DNA hybridization between two complementary domains. Detailed domain design is available in Supporting Information. Each node targets a designated cancer cell membrane receptor via aptamer-receptor binding. When operating the reaction network, we first mix the nodes with the cancer cells in a reaction buffer. If both targeted receptors exist on the membrane, both nodes will be localized on the membrane by aptamer-receptor binding. We then filter out the free nodes in the buffer to exclude potential non-localized reactions. **(b)** DNA strand displacement reactions in the 2-layer linear cascade. The initiator starts the cascade reaction between node A and node B. The output of node B reacts with the reporter complex to tag the cell by a fluorophore. Note that reaction networks that are more complex than linear cascades can be built using the same strategy. **(c)** Abstraction of the 2-layer linear cascade in (b). In the abstraction, we use the name of the aptamer of a node to denote the node. An arrow to a node indicates its input and an arrow from a node indicates its output.

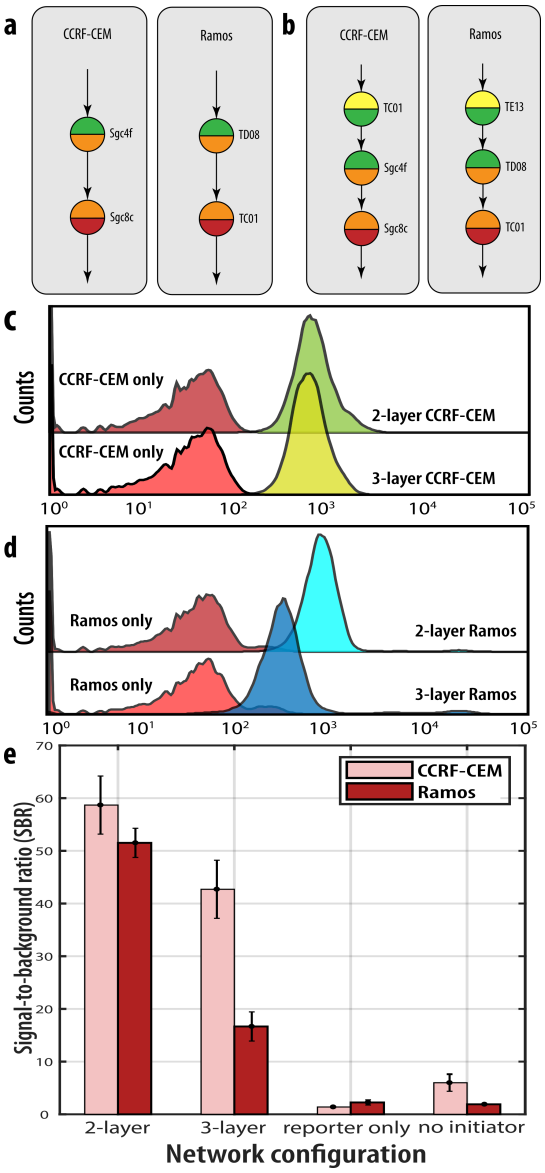

Figure 2. Linear cascades on CCRF-CEM and Ramos. (a), (b) 2-layer and 3-layer linear cascades on CCRF-CEM and Ramos. **(c)** Flow cytometry result of a single repeat of testing 2-layer and 3-layer cascades on CCRF-CEM. Using the 2-layer cascade to explain, the cell population treated by the cascade has much stronger fluorescence intensity than the cell population without any treatment. We get the geometric means of fluorescence intensity for both populations, and calculate the ratio between the two geometric means (green population over red population) to get a signal-to-background ratio (SBR). We repeat such experiment for *three times* to get three SBRs and the statistics in (e). It is the same for all reaction networks demonstrated in this paper. Note that the horizontal axis is fluorescence intensity (log-scale) and the vertical axis is cell count. **(d)** Flow cytometry result of a single repeat of testing 2-layer and 3-layer cascades on Ramos. **(e)** Statistics of SBRs of linear cascades and control experiments.

to describe reaction networks, we introduce an abstraction (Figure 1 (c)) where a node is denoted by the name of its aptamer. If the output of a node is the input of another node, we put an arrow between two nodes.

To demonstrate our architecture, we first implemented 2-layer and 3-layer linear cascades on CCRF-CEM and Ramos cancer cell lines (Figure 2 (a), (b)). All linear cascades worked well and gave strong fluorescence signal to the cells (Figure 2 (c), (d)). We define signal-to-background ratio *(SBR)* as the ratio between the geometric means of fluorescence intensity of the labeled and unlabeled cell populations. We repeat such experiments for three times and get the statistics of SBR in Figure 2 (e). Fluorescence intensity of cells is characterized using flow cytometry. The 3-layer linear cascade on Ramos gives much smaller ratios compared to the 2-layer cascade, and this may be because the density of the receptor for aptamer TE13 is low, or the

two receptors for aptamers TE13 and TD08 do not get close to each other on the cell membrane and then the corresponding nodes cannot interact efficiently.

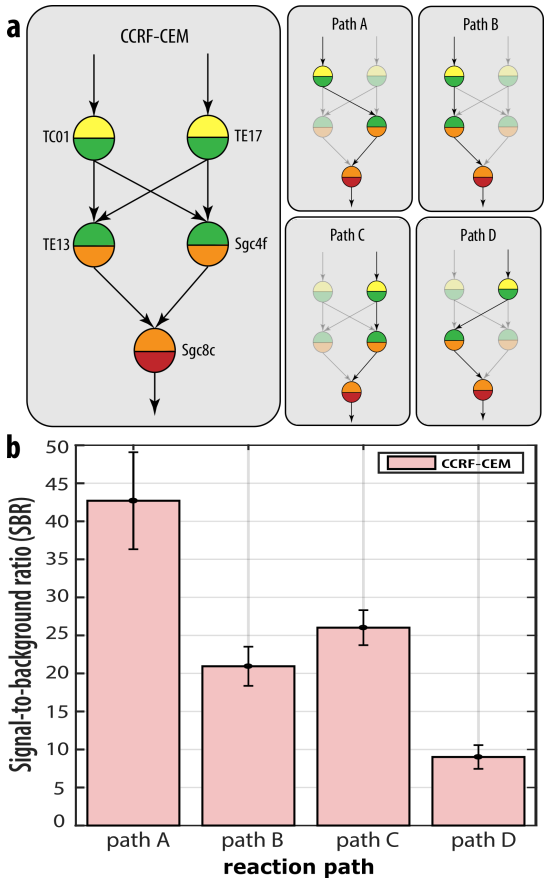

Figure 3. A complex reaction network on CCRF-CEM. (a) A reaction network of five nodes on CCRF-CEM and the four paths to produce output. **(b)** Flow cytometry result of testing the four paths. Statistics of SBRs are from three repeats of each case.

Complex reaction networks beyond linear cascades can also be implemented by our architecture. Figure 3 (a) shows a reaction network on CCRF-CEM that has five nodes. There are four possible linear paths from an initiator to the output of the network. To demonstrate the reaction network, we tested all four paths. The flow cytometry results show that almost all paths give good SBRs (Figure 3 (b)). From the result, we can see that the nodes for TE13 and TE17 are the bottleneck. When neither is in the path (path A), the average SBR is the highest. When one of them is in the path (paths B and C), the average SBR is lower. When both are in the path (path D), the average SBR is the lowest. This may be because the receptors of TE13 and TE17 are at low density on CCRF-CEM.

The reaction networks by our architecture can perform logic computation. For example, the network in Figure 4 (a) can be used as a logic AND gate, where the network only works on a cell membrane having the receptors for both aptamers. We tested this network on four cell lines, and only CCRF-CEM has both receptors. The flow cytometry result shows that only the experiments on CCRF-CEM give large SBRs (Figure 4 (b)). The small non-zero SBRs of other cell lines may be due to random binding of reporter complexes on cell membranes. Therefore, this network can be used to identify CCRF-CEM from other cell lines. The SBR difference

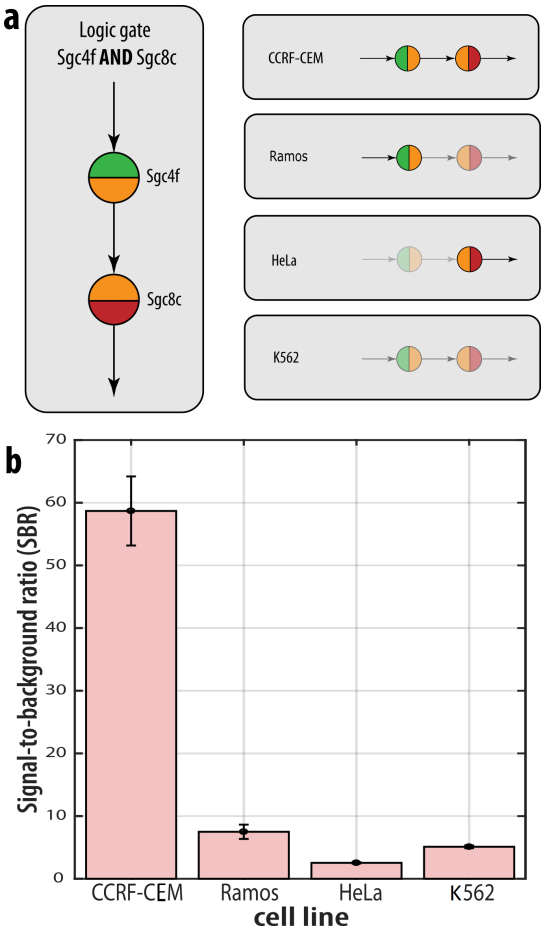

Figure 4. A linear cascade doing logic computation to recognize CCRF-CEM. (a) A linear cascade that can perform logic AND operation, and recognize CCRF-CEM because only CCRF-CEM has the receptors for both aptamers (indicated by solid nodes). **(b)** Flow cytometry result of testing the network in (a) on four cell lines. Statistics of SBRs are from three repeats of each case.

between CCRF-CEM and other cell lines is larger than by prior methods.[8,9] Although HeLa has nonzero (but low) density of the receptor for aptamer Sgc4f,[9] we did not observe relatively high SBRs. This may be because the receptors for aptamers Sgc4f and Sgc8c can hardly reach each other on the membrane of HeLa. More complicated logic computation can also be done. For example, the reaction network in Figure 3 (a) can evaluate the logic function (TC01 OR TE17) AND (Sgc4f OR TE13) AND Sgc8c.

Control experiments were conducted on both CCRF-CEM and Ramos. We mixed a cell population with reporter complex without any reaction network, to see whether it made a difference in SBR. The point here is to make sure that the high fluorescence intensity of a cell population treated by a reaction network in prior experiments is because of the reaction network, not random binding of reporter complex to a cell membrane. As shown in Figure 2 (e), a cell population mixed with only reporter complex does not have a much different fluorescence intensity from a cell population without any treatment, so the SBRs are small. Another control experiment we did was about the leak (reaction without initiator) in the networks, where we mixed a cell population with a network and reporter complex without any initiator,

to see whether it made a difference in SBR. For the case of CCRF-CEM, we added in all nodes in Figure 3 (a) to see the maximum possible leak, where the 2-layer and 3-layer cascades in Figure 2 were already included (in path A). For the case of Ramos, we added in all nodes of the 3-layer cascade in Figure 2 (b) to see the maximum possible leak, where the 2-layer cascade was part of it. As shown in Figure 2 (e), the leak is minimal. These control experiments help to prove that the reaction networks have worked as designed on cell membranes to tag cells with the fluorophore.

Supporting Information

Materials and methods; detailed node design; flow cytometry data; DNA sequences; cell lines and corresponding aptamers used; discussion.

Notes

The authors do not have competing interests.

Acknowledgements

This work is supported by NSF Grants CCF-1320360, CCF-1217457, CCF-1617791 and CCF-1813805.

1

References

(1) Yin, P.; Choi, H. M.; Calvert, C. R.; Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **2008**, *451*, 318–322.
(2) Seelig, G.; Soloveichik, D.; Zhang, D. Y.; Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **2006**, *314*, 1585–1588.
(3) Qian, L.; Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **2011**, *332*, 1196–1201.
(4) Douglas, S. M.; Bachelet, I.; Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **2012**, *335*, 831–834.
(5) Soloveichik, D.; Seelig, G.; Winfree, E. DNA as a universal substrate for chemical kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 5393–5398.
(6) Rudchenko, M.; Taylor, S.; Pallavi, P.; Dechkovskaia, A.; Khan, S.; Butler Jr, V. P.; Rudchenko, S.; Stojanovic, M. N. Autonomous molecular cascades for evaluation of cell surfaces. *Nat. Nanotechnol.* **2013**, *8*, 580–586.
(7) Wu, C.; Han, D.; Chen, T.; Peng, L.; Zhu, G.; You, M.; Qiu, L.; Sefah, K.; Zhang, X.; Tan, W. Building a multifunctional aptamer-based DNA nanoassembly for targeted cancer therapy. *Journal of the American Chemical Society* **2013**, *135*, 18644–18650.
(8) You, M.; Peng, L.; Shao, N.; Zhang, L.; Qiu, L.; Cui, C.; Tan, W. DNA "nano-claw": logic-based autonomous cancer targeting and therapy. *J. Am. Chem. Soc.* **2014**, *136*, 1256–1259.
(9) You, M.; Zhu, G.; Chen, T.; Donovan, M. J.; Tan, W. Programmable and multiparameter DNA-based logic platform for cancer recognition and targeted therapy. *J. Am. Chem. Soc.* **2014**, *137*, 667–674.
(10) Qiu, L.; Zhang, T.; Jiang, J.; Wu, C.; Zhu, G.; You, M.; Chen, X.; Zhang, L.; Cui, C.; Yu, R.; Tan, W. Cell membrane-anchored biosensors for real-time monitoring of the cellular microenvironment. *Journal of the American Chemical Society* **2014**, *136*, 13090–13093.
(11) Han, D.; Wu, C.; You, M.; Zhang, T.; Wan, S.; Chen, T.; Qiu, L.; Zheng, Z.; Liang, H.; Tan, W. A cascade reaction network mimicking the basic functional steps of adaptive immune response. *Nature chemistry* **2015**, *7*, 835.
(12) Yang, Y.; Goetzfried, M. A.; Hidaka, K.; You, M.; Tan, W.; Sugiyama, H.; Endo, M. Direct visualization of walking motions of photocontrolled nanomachine on the DNA nanostructure. *Nano letters* **2015**, *15*, 6672–6676.
(13) Song, T.; Garg, S.; Mokhtar, R.; Bui, H.; Reif, J. Analog computation by DNA strand displacement circuits. *ACS Synth. Biol.* **2016**, *5*, 898–912.
(14) Fern, J.; Scalise, D.; Cangialosi, A.; Howie, D.; Potters, L.; Schulman, R. DNA strand-displacement timer circuits. *ACS Synth. Biol.* **2017**, *6*, 190–193.
(15) He, L.; Lu, D.; Liang, H.; Xie, S.; Zhang, X.; Liu, Q.; Yuan, Q.; Tan, W. mRNA-Initiated, Three-Dimensional DNA Amplifier Able to Function inside Living Cells. *Journal of the American Chemical Society* **2017**, *140*, 258–263.
(16) Zhao, B.; O'Brien, C.; Mudiyanselage, A. P. K.; Li, N.; Bagheri, Y.; Wu, R.; Sun, Y.; You, M. Visualizing intercellular tensile forces by DNA-based membrane molecular probes. *J. Am. Chem. Soc.* **2017**, *139*, 18182–18185.
(17) You, M.; Lyu, Y.; Han, D.; Qiu, L.; Liu, Q.; Chen, T.; Wu, C. S.; Peng, L.; Zhang, L.; Bao, G.; Tan, W. DNA probes for monitoring dynamic and transient molecular encounters on live cell membranes. *Nat. Nanotechnol.* **2017**, *12*, 453.
(18) Peng, R.; Wang, H.; Lyu, Y.; Xu, L.; Liu, H.; Kuai, H.; Liu, Q.; Tan, W. Facile assembly/disassembly of DNA nanostructures anchored on cell-mimicking giant vesicles. *Journal of the American Chemical Society* **2017**, *139*, 12410–12413.
(19) Wan, S.; Zhang, L.; Wang, S.; Liu, Y.; Wu, C.; Cui, C.; Sun, H.; Shi, M.; Jiang, Y.; Li, L.; Qiu, L.; Tan, W. Molecular recognition-based DNA nanoassemblies on the surfaces of nanosized exosomes. *Journal of the American Chemical Society* **2017**, *139*, 5289–5292.
(20) Thubagere, A. J.; Li, W.; Johnson, R. F.; Chen, Z.; Doroudi, S.; Lee, Y. L.; Izatt, G.; Wittman, S.; Srinivas, N.; Woods, D.; Winfree, E.; Qian, L. A cargo-sorting DNA robot. *Science* **2017**, *357*, eaan6558.
(21) Cangialosi, A.; Yoon, C.; Liu, J.; Huang, Q.; Guo, J.; Nguyen, T. D.; Gracias, D. H.; Schulman, R. DNA sequence–directed shape change of photopatterned hydrogels via high-degree swelling. *Science* **2017**, *357*, 1126–1130.
(22) Srinivas, N.; Parkin, J.; Seelig, G.; Winfree, E.; Soloveichik, D. Enzyme-free nucleic acid dynamical systems. *Science* **2017**, *358*, eaal2052.
(23) Song, T.; Garg, S.; Mokhtar, R.; Bui, H.; Reif, J. Design and analysis of compact DNA strand displacement circuits for analog computation using autocatalytic amplifiers. *ACS Synth. Biol.* **2018**, *7*, 46–53.
(24) Cherry, K. M.; Qian, L. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature* **2018**, *559*, 370.
(25) Kishi, J. Y.; Schaus, T. E.; Gopalkrishnan, N.; Xuan, F.; Yin, P. Programmable autonomous synthesis of single-stranded DNA. *Nat. Chem.* **2018**, *10*, 155.
(26) Fern, J.; Schulman, R. Modular DNA strand-displacement controllers for directing material expansion. *Nat. Commun.* **2018**, *9*, 3766.
(27) Peng, R.; Zheng, X.; Lyu, Y.; Xu, L.; Zhang, X.; Ke, G.; Liu, Q.; You, C.; Huan, S.; Tan, W. Engineering a 3D DNA-logic gate nanomachine for bispecific recognition and computing on target cell surfaces. *Journal of the American Chemical Society* **2018**, *140*, 9793–9796.
(28) Liu, H.; Yang, Q.; Peng, R.; Kuai, H.; Lyu, Y.; Pan, X.; Liu, Q.; Tan, W. Artificial Signal Feedback Network Mimicking Cellular Adaptivity. *Journal of the American Chemical Society* **2019**, *141*, 6458–6461.
(29) Xiong, M.; Rong, Q.; Kong, G.; Yang, C.; Zhao, Y.; Qu, F.-L.; Zhang, X.-B.; Tan, W. Hybridization chain reaction-based nanoprobe for cancer cell recognition and amplified photodynamic therapy. *Chemical Communications* **2019**, *55*, 3065–3068.
(30) Chang, X.; Zhang, C.; Lv, C.; Sun, Y.; Zhang, M.; Zhao, Y.; Yang, L.; Han, D.; Tan, W. Construction of a Multiple-Aptamer-Based DNA Logic Device on Live Cell Membranes via Associative Toehold Activation for Accurate Cancer Cell Identification. *Journal of the American Chemical Society* **2019**, *141*, 12738–12743.
(31) Shangguan, D.; Li, Y.; Tang, Z.; Cao, Z. C.; Chen, H. W.; Mallikaratchy, P.; Sefah, K.; Yang, C. J.; Tan, W. Aptamers evolved from live cells as effective molecular probes for cancer study. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 11838–11843.
(32) Fang, X.; Tan, W. Aptamers generated from cell-SELEX for molecular medicine: a chemical biology approach. *Acc. Chem. Res.* **2009**, *43*, 48–57.
(33) Breaker, R. R. Natural and engineered nucleic acids as tools to explore biology. *Nature* **2004**, *432*, 838.
(34) Sefah, K.; Shangguan, D.; Xiong, X.; O'donoghue, M. B.; Tan, W. Development of DNA aptamers using Cell-SELEX. *Nat. Protoc.* **2010**, *5*, 1169.
(35) Famulok, M.; Hartig, J. S.; Mayer, G. Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy. *Chem. Rev.* **2007**, *107*, 3715–3743.

Graphical TOC Entry