

MOLECULAR ENGINEERING OF DNA-BASED SYSTEMS FOR INTELLIGENT  
THERAPY

By

DA HAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2013

© 2013 Da Han

To my family and friends

## ACKNOWLEDGMENTS

It has been an amazing journey for me as a graduate student in Tan group in the past 5 years. This journey was full of happiness, determination, sweat and difficulties, which I define it as “where the dream starts”. I owe many thanks to a lot of people who I will remember forever.

First, I want to express my most tremendous gratitude to my advisor, Dr. Weihong Tan. As a student of his, I was given a lot of freedom to do the research that I had interest in and comforted with great encouragement when confronting difficulties. He teaches me a lot on how to think, read and write, which have improved my performance as a chemist. Meanwhile, the way he treats other people and life has also benefited and inspired me a lot. I believe those experiences will be a great fortune in my future life.

I also want to thank my committee members Dr. Nicolas C. Polfer, Dr. Rebecca A. Butcher, Dr. Y. Charles Cao, and Dr. Gregory Schultz for their valuable advice, support, and encouragement. In addition, I also express my gratitude to Dr. Kathryn R. Williams for her help on manuscript writing and review.

Only with the great support from the people in Tan group, this dissertation can be successfully completed. I want to thank Dr. Yan Chen, Dr. Jin Huang, Dr. Zhi Zhu and Dr. Quan Yuan for their guidance on involving me into my research. I also appreciate the help from Dr. Kwame Sefah, Dr. Tao Zhang, Dr. Xiangling Xiong, Dr. Ibrahim Shukoor, Dr. Basri Gulbakan, Dr. Tao Chen, Dr. Lu Peng, Dr. Guizhi Zhu, Dr. Elizabeth Jimenez and Dr. Mingxu You on scientific discussion and instrumental training. Also, I thank the useful and interesting suggestions from other Tan group members including Ismail Ocsoy, Yunfei Zhang, Cuichen Wu, Sena Cansiz, Diane Turek, Emir Yasun, Liqin

Zhang, Cheng Cui, Liping Qiu and Carole Champanhac. Each of them has made my journey enjoyable and pleasant.

Finally, I deeply indebted my parents and would like to express my greatest respect to them. Without their support, encouragement and love, I cannot complete anything. So I want to dedicate this dissertation and show my love to them; my heart always goes with them.

## TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF TABLES.....	9
LIST OF FIGURES.....	10
ABSTRACT .....	13
CHAPTER	
1 INTRODUCTION .....	16
Nucleic Acids and the “Art” of Molecular Engineering.....	16
Structures and Chemical Synthesis of Nucleic Acids .....	17
Functional Nucleic Acids: Aptamers .....	19
Photocontrollable DNA Nanostructures .....	21
DNA Nanostructures .....	21
Azobenzene and Azobenzene Incorporated DNA .....	23
DNA-Based Logical Devices.....	24
Overview of DNA-based Computation.....	24
Construction of DNA Circuit: Toehold-Mediated Strand Displacement.....	25
DNA-based Circuit for Analytical and Biomedical Applications.....	27
DNA-based Biomimicry .....	28
2 MOLECULAR ENGINEERING OF A CELL-SURFACE APTAMER CIRCUIT FOR TARGETED AND AMPLIFIED PHOTODYNAMIC CANCER THERAPY .....	39
Significance and Background .....	39
Experimental Materials and Methods.....	43
Cell Culture.....	43
Ce6-Modified DNA Synthesis .....	43
DNA Purification .....	44
Fluorescence Kinetics of DNA Hairpin Circuit in Buffer .....	44
Test of Fluorescence Response and SOG of DNA Hairpin Circuit .....	45
Test of SOG Response of DNA Hairpin Circuit to Cancer Cells .....	46
Flow Cytometric Analysis .....	46
Confocal Imaging of Cells Bound with Aptamer .....	47
Cytotoxicity Study.....	47
Results and Discussion.....	48
Signal Amplification Effect of Aptamer Circuit in Buffer .....	48
Selective Recognition Ability of the Aptamer Circuit.....	50
Singlet Oxygen Amount with Different Cancer Cells .....	51
Selective Cytotoxicity Effect to Cancer Cells .....	51
Conclusions .....	52

3	MOLECULAR ENGINEERING OF PHOTORESPONSIVE THREE-DIMENSIONAL DNA NANOSTRUCTURES FOR INTELLIGENT DRUG DELIVERY SIGNIFICANCE AND BACKGROUND .....	61
	Experimental Materials and Methods.....	63
	Chemicals and Regents .....	63
	Synthesis of DNA Sequences .....	63
	Synthesis of BF-PS-DNA.....	64
	Assembly of DNA Tetrahedral Nanostructures.....	64
	Native Polyacrylamide Gel Electrophoresis (PAGE) for Structural Characterization .....	65
	Phosphorothioate DNA Tetrahedron Assembled with Gold Nanoparticles .....	65
	FRET Measurement of Structure Change in Response to UV and Visible Irradiation .....	66
	AFM and TEM Measurements.....	66
	Results and Discussion.....	67
	Construction and Optimization of DNA Tetrahedral Nanostructure .....	67
	Structural Confirmation of DNA Tetrahedral Nanostructure.....	67
	Characterization of DNA tetrahedral structure with photocontrollability .....	68
	Conclusions .....	70
4	MOLECULAR ENGINEERING OF A DNA LOGICAL CIRCUIT FOR PROGRAMMABLE AND AUTONOMOUS REGULATION OF PROTEIN ACTIVITY.....	76
	Significance and Background .....	76
	Experimental Materials and Methods.....	81
	DNA Synthesis .....	81
	DNA Purification .....	82
	Quantitative Analysis of Released DNA-input Strands Triggered by Thrombin .....	82
	Validation of Signal Transduction by Fluorescence .....	83
	Thrombin Catalytic Activity Assay .....	84
	Results and Discussion.....	85
	DNA Aptamer Circuit Design .....	85
	Validation of Signal Transduction in Each Module .....	86
	Performance of Entire Circuit with Thrombin as Input .....	88
	Programmable and Autonomous Protein Regulation by Aptamer Circuit .....	88
	Theoretical calculations of circuit performance.....	92
	Conclusions .....	97
5	MOLECULAR ENGINEERING OF A DNA-ENZYME CASCADE NETWORK FOR ACQUIRED IMMUNE SYSTEM MIMICRY.....	112
	Significance and Background .....	112
	Experimental Materials and Methods.....	113
	DNA Synthesis .....	113

DNA Purification .....	114
Preparation of Circular Template DNA (CP).....	114
Validation of Signal Transduction in Each step and the Entire System by Fluorescence.....	115
Results and Discussion.....	115
System Design and Construction .....	115
Validation of Signal Transduction in Each Module .....	119
Performance of the Entire System with Pathogen Input .....	123
Conclusions .....	125
6 FUTURE DIRECTIONS AND CONCLUSIONS .....	136
Conclusions .....	136
Future Directions .....	139
LIST OF REFERENCES .....	141
BIOGRAPHICAL SKETCH.....	153

## LIST OF TABLES

<u>Table</u>		<u>page</u>
2-1	Sequence of oligonucleotides used in this chapter. LNA bases are indicated by bold and underscores. ....	54
3-1	DNA sequences used for tetrahedron assembly .....	72
4-1	Sequences of molecular circuit for spectrofluorimetry studies.....	99
4-2	Sequences of molecular circuit for chromogenic peptide substrate hydrolysis monitoring.....	100
5-1	Sequences used in the AIS mimicry system.....	126

## LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 Chemical structure of nucleic acids and nucleosides. ....	30
1-2 Structure of DNA duplex and scheme of Chargaff's rule. ....	31
1-3 Chemical structure of DNA phosphoramidite.....	32
1-4 Automated oligonucleotide synthesis based on phosphoramidite chemistry. ....	33
1-5 Scheme of Systematic Evolution of Ligands by EXponential enrichment (SELEX). ....	34
1-6 A) Illustration of some DNA tiles. <sup>61</sup> B) Schematic folding of DNA origami.. ....	35
1-7 Scheme of the reversible hybridization/dehybridization of an azobenzene incorporated DNA duplex. ....	36
1-8 A) Scheme of toehold-mediated displacement reaction. B) Scheme of toehold-exchange displacement reaction. ....	37
1-9 Example of functional nucleic acid-based circuit for biological and biomedical applications.....	38
2-1 Illustration of DNA aptamer circuit on cell membrane.....	55
2-2 Image of the PAGE gel proving the catalytic effect of C sequence.....	56
2-3 Signal amplification effect of the aptamer circuit in buffer.....	57
2-4 Fluorescence kinetics describing the leakage reaction of $A_1$ , $A_2$ and $R_{12}$ in the Fluo buffer.. ....	58
2-5 Fluorescence spectra of mixtures containing 100 nM $A_1$ , 100 nM $A_2$ and 150 nM Ce6-modified $R_{12}$ with different concentrations of TDO5-C in buffer. ....	58
2-6 Demonstration of the selective recognition ability of the aptamer circuit.. ....	59
2-7 SOSG fluorescence of DNA circuit ( $A_1$ , $A_2$ and $R_{12}$ ) incubated with buffer, TDO5-C-labeled CCEF-CEM cells (control) and TDO5-C-labeled Ramos cells (target).....	60
2-8 Cell viability result using MTS assay.. ....	60
3-1 Design of photocontrollable DNA nanostructure using azobenzenes.....	73

3-2	A) Native polyacrylamide gel electrophoresis (6%) analysis of different azo-incorporated tetrahedral structures at 4 °C. B) AFM image of T-10Azo structures recorded with 10nm tips.....	74
3-3	Fluorescent characterization of DNA tetrahedral structure with photocontrollability.....	74
3-4	Illustration of gold nanoparticle (AuNP) assembly on three vertices of the variational triangular faces of tetrahedra by a phosphorothioate anchor and a short bifunctional fastener (BF).....	75
3-5	TEM images of AuNPs assembled on three vertices of the variational triangular face of tetrahedra..	75
4-1	Working Scheme of molecular circuit..	101
4-2	Reaction pathways of DNA-input and Output (O) with the aid of Fuel (F). The design follows the procedure described by Qian and Winfree. <sup>51</sup>	102
4-3	Absorbance of chromogenic peptide hydrolysis product <b>vs.</b> time for different concentrations of thrombin..	103
4-4	A) Absorbance change of chromogenic hydrolysis product as a function of time with 300 nM thrombin preincubated with 150 nM, 300 nM, 600 nM and 1.5 μM TA-15. B) Absorbance change of chromogenic hydrolysis product as a function of time with 300 nM thrombin preincubated with 1.5 μM TA-15 and TA-29.....	103
4-5	Absorbance of chromogenic peptide hydrolysis product as a function of time in the presence of different thrombin concentrations with the aptamer circuit..	104
4-6	Fluorescence kinetics of FRET-based A-I duplex (100 nM) in the presence of 0 nM, 200 nM, and 500 nM thrombin.....	104
4-7	Validation of the signal transduction in the Input Convertor module.....	105
4-8	Fluorescence calibration curve for different concentrations of FAM-labeled TA-29.....	105
4-9	Validation of the signal transduction in the Threshold Controller and Inhibitor Generator modules.....	106
4-10	Fluorescence output versus DNA-input concentration plot.....	107
4-11	Fluorescence output versus thrombin concentration plot with two preset threshold values (concentration of Tmb) at 100 nM and 200 nM.....	108
4-12	Inhibitory effects of different concentrations of TA-15 and TA-29 on thrombin.	109

4-13	Absorbance changes of chromogenic peptide substrate for different thrombin concentrations using the design in Figure 4-1. ....	109
4-14	Working scheme of Inhibitor Generator 2. ....	110
4-15	A) Absorbance changes ( $\delta A$ ) of chromogenic peptide substrate upon different thrombin concentrations with and without circuit. B) $V_{obs}$ of the enzymatic reaction upon different thrombin concentrations with and without circuit. ....	110
4-16	A). Plot of theoretical [DNA-input] vs. time. B). Plot of theoretical [Tmb-TA15] vs. time. ....	111
4-17	A) Plot of [DNA-input] vs. $K_{eq}$ for different $a$ values, where $a$ is the concentration ratio of [Tmb] and [A-I]. B) Plot of $\Delta$ [DNA-input] vs. $K_{eq}$ . Here, $\Delta$ [DNA-input] was calculated by using $f$ ([DNA-input] <sub><math>a=4</math></sub> ) – $f$ ([DNA-input] <sub><math>a=0.5</math></sub> ). C) The zoom-in plot of (B). ....	111
5-1	Working principle of AIS biomimicry system. ....	127
5-2	The proposed entropy-driven catalytic pathway. <sup>49</sup> ....	128
5-3	Fluorescence studies of the reaction priority between steps 1 and 2, as well as the catalytic effect of TM on the reaction of BM and P <sub>0</sub> . ....	129
5-4	Fluorescence kinetics of the system with different P <sub>0</sub> additions. ....	129
5-5	Analysis by PAGE (10% native gel) of the reaction pathway in step 1 and 2. .	130
5-6	A) Scheme of primer-assisted RCA reaction and plot of fluorescence restoration versus different concentrations of primer 12a. B) Scheme of the design of mismatch points on PG and result of fluorescence enhancement fold versus different concentrations of ssDNA PI and Phi29. ....	131
5-7	Scheme of the primer-initiated RCA product cut by SsPI restriction enzyme and their analysis by PAGE (8% denatured gel at 4 °C). ....	132
5-8	Experimental results of the entire system with pathogen input. ....	133
5-9	Gel analysis by PAGE (8% denatured gel) for the entire AIS mimicry system with SsPI restriction enzyme. ....	134
5-10	Analysis by PAGE (8% denatured gel) of the preparation of circular template (CP). ....	135

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

MOLECULAR ENGINEERING OF DNA-BASED SYSTEMS FOR INTELLIGENT  
THERAPY

By

Da Han

December 2013

Chair: Weihong Tan  
Major: Chemistry

Nucleic acids, as biological macromolecules essential for life, play major roles in transferring, encoding and expressing genetic information across generations, as well as in regulating biological reactions. As an effective way to explore unique and versatile functions of nucleic acids, molecular engineering focuses on understanding the physical behavior of nucleic acids under different conditions and the effects of possible chemical modifications. The central theme of this thesis is to develop “smart” DNA systems for intelligent therapy on the molecular level via molecular engineering methods.

In the first phase of the research, an aptamer-based DNA nanocircuit was designed for the selective recognition of cancer cells, controllable activation of a photosensitizer and amplification of the photodynamic therapeutic effect. The aptamers can selectively recognize target cancer cells and bind to the specific proteins on the cell membrane. Then the overhanging catalyst sequence on the aptamer triggers a toehold-mediated catalytic strand displacement to activate the photosensitizer and produce reactive singlet oxygen. The specific binding-induced activation allows the DNA circuit to distinguish diseased cells from healthy cells, reducing damage to nearby healthy cells. Moreover, the catalytic amplification reaction takes place close to the target

cancer cells, resulting in a high local concentration of singlet oxygen to selectively kill the target cells. The principle employed in this study demonstrated the feasibility of assembling a DNA circuit on cell membranes and could further broaden the utility of DNA circuits for applications in biology, biotechnology, and biomedicine.

In the second project, azobenzene-incorporated DNA was used as a control agent to precisely monitor three-dimensional DNA nanostructures. The shape of a DNA tetrahedron can be controlled by alternate irradiations with different wavelengths of light, thus enabling the future application for intelligent drug delivery manipulated by photons.

The third project involved describing the first logical circuit based on DNA-protein interactions with accurate threshold control, enabling autonomous, self-sustained and programmable manipulation of protein activity *in vitro*. Previous designs were based primarily on DNA hybridization and strand displacement reactions. The new design uses diverse nucleic acid interactions with proteins. The circuit can precisely sense the local enzymatic environment, such as the concentration of thrombin, and when it is excessively high, a coagulation inhibitor is automatically released by a concentration-adjusted circuit module. To demonstrate the programmable and autonomous modulation, a molecular circuit with different threshold concentrations of thrombin was tested as a proof of principle. In the future, the advantages of tunable regulation, design modularity and target specificity of this prototype could lead to the development of novel DNA biochemical circuits to control the delivery of aptamer-based drugs in smart and personalized medicine, providing a more efficient and safer therapeutic strategy.

Finally, an artificial DNA biomimetic network was developed for vertebrate acquired immune system. The dynamic networks of biochemical reactions display

extremely complex functional behaviors which can be classified as specific recognition, antigen tolerance, immune response, and long-term memory. We have reproduced such functions by the rational design of dynamic reaction networks based on DNA biochemistry. Rather than enrolling cellular components as in the real acquired immune system *in vivo*, we used DNA and enzymes as simplified and artificial analogs to mimic their systematic response to specific molecular stimuli *in vitro*. Building on the successful implementation of each step, our results show patterns that follow the macroscopic behaviors of vertebrate acquired immune system. This approach provides rational and simplified bottom-up construction strategies to design complex biomimetic reaction networks.

## CHAPTER 1 INTRODUCTION

### **Nucleic Acids and the “Art” of Molecular Engineering**

Battling diseases has been a never-ending task in human history. Current studies reveal that many diseases are characterized by abnormal behavior on the molecular level, such as deviant gene expression. The key to the effective and successful treatment of disease is the development of intelligent diagnostic and therapeutic methods, including specific recognition of targets, precise control of drug release with temporal and spatial resolution, and effective inhibition of drug side effects. Molecular tools that can specifically recognize disease-related biomarkers, and then efficiently regulate abnormal functions of targets will greatly facilitate the development of intelligent therapy. Among many potential molecular tools, functional nucleic acid tools hold great promise in a variety of diagnostic and therapeutic applications.

Functional nucleic acid probes, such as molecular beacon biosensors,<sup>1, 2</sup> and molecular targeting ligands,<sup>3</sup> have shown to be sensitive and selective in recognizing disease-related biomarkers. In addition, nucleic acid nanotechnology has been developed by employing DNA and RNA as basic building blocks for designing various nanostructures and nanocircuits which may be used for material, biological and medical applications.<sup>4, 5</sup> The investigation of the physical behaviors of nucleic acids under different conditions, as well as their possible chemical properties, are critical factors in the development of new functions. Consequently, these studies have given rise to a new subject area called “nucleic acid-based molecular engineering”. This dissertation focuses on engineering nucleic acids into molecular tools for intelligent diagnosis and therapy of diseases.

## Structures and Chemical Synthesis of Nucleic Acids

Nucleic acids are biological macromolecules essential for life, including DNA (deoxyribonucleic acid) and RNA (ribonucleic acid). The basic building blocks of nucleic acids are called nucleotides, comprised of a pentose sugar ring (2-deoxyribose in DNA and ribose in RNA), a phosphate and a nucleobase (purine or pyrimidine). Different nucleotides are linked to each other in an alternating chain to form an oligonucleotide (sugar-phosphate backbone) through phosphodiester linkages. (Fig.1-1)

The topological structure of DNA was first discovered by James Watson and Francis Crick in 1953.<sup>6</sup> In their models, two single-stranded oligonucleotides form an anti-parallel duplex structure following the base complementary principle (Chargaff's rule), in which specific bases interact with each other mainly through hydrogen bonds. (Fig.1-2) There are five types of natural bases to compose DNA and RNA sequences: adenine (abbreviated A), cytosine (C), guanine (G), thymine (T) only found in DNA, and uracil (U) only found in RNA. According to Chargaff's rule, for DNA duplexes the hydrogen bond pairs are between G-C and A-T, while for RNA duplexes, the pairs are G-C and A-U. Noticeably, the sequence stability of DNA duplex is dependent on the DNA sequence length and the number and ratio of G-C base pairs, which have three hydrogen bonds, to A-T pairs (two hydrogen bonds).

As important biological macromolecules, nucleic acids mainly carry the functions of transferring and encoding genetic information across generations. Furthermore, they are also extremely important for regulation of biological reactions, as well as expression of numerous functional components.<sup>7</sup> Many advances in nucleic acid chemistry have advanced since 1992, when Beaucage developed an automated solid-support system for synthesis of nucleic acid chains via phosphoramidite chemistry.<sup>8,9</sup> As shown in

Fig.1-3, the phosphoramidites are composed of a base, a sugar ring and protection groups (dimethoxytrityl (DMT), diisopropylamino, and 2-cyanoethyl). In addition, a solid-support called CPG (controlled pore glass) is used to load the starting reactant phosphoramidites. The basic procedure includes: (1) dichloroacetic acid (DCA) detritylation of DMT (dimethoxytrityl)-protected phosphoramidite to form the 5'-hydroxyl group; (2) tetrazole activation to protonate the nitrogen of the incoming nucleotide and to form the phosphate linkage; (3) acetic anhydride and N-methylimidazole capping to block the remaining active 5'-hydroxyl group; and finally (4) iodine oxidation to generate the phosphate ester bond. (Fig.1-4) Other modifications, such as addition of fluorophores, quenchers, biotin etc, can be introduced into any desired position of an oligonucleotide, if they can be converted into a phosphoramidite derivative and is compatible with DNA synthesis process. To facilitate the synthesis yield and speed, nucleic acid synthesizer is mainly applied to implement the synthesis procedure.

After many cycles of this synthesis, oligonucleotides with pre-designed sequences and lengths can be prepared. However, due to incomplete reactions and limited chemical reaction yields, the products actually contain a mixture of multiple oligonucleotide species in different lengths requiring off-machine purification, which is usually carried out using high-performance liquid chromatography (HPLC). Specifically, the nucleic acids tethered on the CPG are cleaved by reaction with weakly basic conditions (generally ammonium hydroxide or a mixture with other additives) at a high temperature. The collected products can then be purified by HPLC with a typical mobile phase of acetonitrile and 0.1M triethylammonium (TEAA) water solution and a reversed phase C-18 column. The purified nucleic acid samples are incubated with weak acid

(generally 80% acetic acid) to remove the DMT group, and the concentrations are determined by measuring the absorbance at 260nm.

### **Functional Nucleic Acids: Aptamers**

As discussed above, both DNA and RNA participate in genetic storage and transfer. Moreover, they can also be employed as enzymes (for catalysis) and receptors (for ligand binding). Increasingly, researchers are making interesting use of these molecules, now collectively called functional nucleic acids. Natural functional nucleic acids include ribozymes, allosteric ribozymes and riboswitches, while artificial functional nucleic acids comprise aptamers, ribozymes and deoxyribozymes by *in vitro* selection. Among the most intensively studied functional nucleic acids are aptamers, single-stranded oligonucleotides that fold into unique three-dimensional structures, allowing them to bind specifically to a broad range of target molecules or even whole cells. They are isolated via an *in vitro* selection process named *systematic evolution of ligands by exponential enrichment* (SELEX) against various targets.<sup>10-13</sup> As mimics of antibodies, aptamers have some unique advantages such as high recognition specificity, excellent stability, reproducible synthesis and non-toxicity, which make them novel molecular probes capable of replacing antibodies. Furthermore, the flexibility and programmability of aptamers provide researchers broad opportunities for molecular engineering, making aptamers adaptive for more innovative applications, such as drug delivery,<sup>14</sup> signal transduction<sup>15</sup> and biosensing.<sup>16</sup>

In the SELEX process, a target and perhaps a possibly negative control are chosen firstly. Then, a DNA/RNA library is designed to have a random sequence of 30-40 bases and two primer sequences flanked by 18-20 bases for polymerase chain

reaction amplification (PCR). Because the DNA/RNA library contains  $10^{13}$ - $10^{16}$  different DNA molecules, suitable aptamers will ultimately be selected from it.

A typical selection cycle is shown in Figure 1-5. First, the library pool is incubated with the target, leading to the binding between the target molecules and specific portions of the library. After physically separating the strong binders from the remains of the pool, the library can be further narrowed using only the high affinity sequences, which is the key concept of the entire SELEX process. As another important step, PCR is used to amplify the remaining eluted sequences after separation from the target.<sup>17</sup> Afterwards, the eluted ssDNA is considered as the enriched pool for the first round and used as library for the next round.

To improve the selectivity of aptamers, counter selection is usually implemented by incubating the amplified sequences with the negative controls. In counter selection, the negative control is incubated with eluted sequences and only the unbound probes are collected and amplified by PCR. In this way, the nucleic acid sequences with unspecific binding can be removed. Herein, the concentration of targets, DNA, ionic strength, incubation time and temperature can be optimized to provide more stringent conditions to favor the selection of aptamers with high affinity. In most cases, the pool can be highly enriched after 10-20 rounds of selection cycles. After sequencing the selected aptamers, it is possible to re-synthesize them and evaluate their binding affinities, i.e.  $K_d$ , which are in the  $\mu\text{M}$  to nM range for useful aptamers.

As a successful expansion of the SELEX technique, our group has developed a whole-cell-SELEX (Cell-SELEX) strategy to generate panels of aptamers recognizing different types of cancer cells, including acute lymphoblastic leukemia (ALL) (T cell

leukemia),<sup>18</sup> liver cancer,<sup>19</sup> acute myeloid leukemia (AML) leukemia,<sup>20</sup> lung cancer,<sup>21, 22</sup> ovarian cancer,<sup>23</sup> B cell lymphoma,<sup>24</sup> colorectal cancer,<sup>25</sup> and breast cancer<sup>26</sup>. Cell-SELEX provides a unique set of capabilities, such as obtaining aptamers that target specific cells without prior knowledge of the molecular differences between target and nontarget cells as well as simultaneous generation of a panel of aptamers, which may have different molecular targets. Cell-SELEX has not only provided specific molecular probes for cancer diagnosis, but has also facilitated clinical therapy to identify new cancer drugs and tumor treatment.

### **Photocontrollable DNA Nanostructures**

#### **DNA Nanostructures**

The DNA molecule has appealing features for use in the field of nano-assembly. First, its minuscule size, with a diameter of about 2 nanometers, matches the size requirement for nano-assembly. Second, the reversible, specific and stable hybridization between complementary strands enables the construction of robust building blocks with various pre-designed shapes. In addition, the possibility of predicting the hybridization behavior by computer programs facilitates programmable and autonomous structural design, thus greatly improving the assembly yield.

The field of DNA nanostructures was pioneered in the 1980s by Nadrian Seeman. In his design, the basic building block is called a DNA tile, which is a DNA nanostructure having a number of sticky ends (pads) on its sides. (Fig.1-6A) Another of his structural concept is termed as DNA lattice, which is a DNA nanostructure composed of a group of DNA tiles that are assembled together via hybridization of their pads. Generally the strands composing the DNA tiles are designed to have melting temperatures above those of the pads, ensuring that when the component DNA

molecules are combined in solution, the DNA tiles assemble first, and only then, as the solution is further cooled, do the tiles bind together via hybridization of their pads. Subsequently, with proper sticky-ends design, these tiles were successfully self-assembled into linear arrays,<sup>27</sup> 2D lattices,<sup>28, 29</sup> and DNA tubes.<sup>30</sup> Figure 1-5 shows some examples of different DNA nanostructures using DNA tiles as building blocks.

Another significant method to prepare DNA nanostructures was developed by Rothemund in 2006.<sup>31</sup> His approach termed “DNA origami”, makes use of a long strand of “scaffold” ssDNA (e.g., from the genome of a viral phage) that has only weak secondary structure. (Fig. 1-6B) After the addition of a large number of relatively short “staple” ssDNA sequences, with subsequences complementary to certain subsequences of the scaffold ssDNA, the long scaffold can be induced to fold into a fully addressable 2D DNA nanostructure. This method allows for the construction of arbitrary patterns with computer assistance, including rectangles, stars, smiley faces, maps and other 2D shapes. In addition, DNA origami provides the opportunity to position molecules or objects with nanoscale precision, for use in nanoparticle assembly<sup>32</sup> and protein nanoarrays.<sup>33, 34</sup>

Other than 2D DNA nanostructures, 3D DNA nanostructures have also been investigated to provide more profound knowledge of the molecular structures of DNA. Shih et al.<sup>35</sup> first built a rigid DNA octahedron by folding a 1.7-kb ssDNA in the presence of five 40-mer helper strands. Turberfield et al. and Mao et al. developed a series of DNA 3D structures, including tetrahedra, dodecahedra, and buckyballs, using a one-pot annealing strategy.<sup>36, 37</sup> Other 3D DNA nanostructures with controllable shape or pattern change include a nanobox<sup>38</sup> with controllable lids and a nanobucket with

intelligent covers.<sup>39</sup> These 3D DNA reconfigurable nanostructures, although still in their infancy, have great potential for use in drug delivery and tissue engineering.

### **Azobenzene and Azobenzene-Incorporated DNA**

Azobenzene is a chemical compound with two phenyl rings linked by a N=N double bond. The most attractive feature of this molecule is photo-isomerization, which can be induced by alternative exposure to light irradiation with different energies (wavelength).<sup>40</sup> Specifically, the azobenzene isomerizes from its planar *trans*- form to the non-planar *cis*- form after UV-light irradiation (300nm ~ 400nm), and back to the *trans*- form after irradiation with visible light (400nm ~ 500nm). This process is completely reversible under UV and visible irradiation, and can be used in triggering photo-initiating or photo-regulating processes.

The successful synthesis of an azobenzene phosphoramidite derivative has enabled the manufacture of azobenzene-incorporated DNA molecules as well as photosensitive DNA nanostructures.<sup>41</sup> The mechanism of this photo-controlled DNA behavior has been shown to result from the isomerization of azobenzene moieties, which induces hybridization/dehybridization between complementary stands.<sup>42</sup> When the azobenzene incorporated DNA is irradiated with visible light, the azobenzene moieties adopt the planar *trans* configuration and stabilize the hybridization by aromatic  $\pi$ - $\pi$  stacking interactions. When irradiated with UV light, the azobenzene moieties isomerize to the non-planar *cis* configuration and disrupt the duplex structure by steric hindrance (Fig.1-6). Therefore, by incorporating azobenzene moieties into DNA nanostructures, shapes mediated by strand hybridization can be controlled by the

interconversion of azobenzene moieties, allowing the construction of photocontrollable DNA nanostructures.<sup>43, 44</sup>

## **DNA-Based Logical Devices**

### **Overview of DNA-based Computation**

Silicon-based computers have been intensively applied in a vast array of areas, because of the properties such as unprecedented computing power, seamless coupling ability and incredible adaptability. However, the development of more powerful microprocessors is heading towards a barrier, due to the physical limitations of conventional silicon chips. Continued progress will require miniaturization their components, which may affect the instrumental performance.<sup>45</sup> To address the challenge ahead, researchers have been pursuing the idea of constructing computers in which computations are performed by individual molecules, as they may allow the continuously exponential increase in performance and decrease in size for microprocessors.

Nucleic acids have been found extremely effective for miniaturization of information storage, because only approximately 50 atoms are used for one bit of information. In addition, the easy chemical synthesis, combinatorial structures, and Watson-Crick complementarity principle provide sufficient theoretical and experimental basis for design of logical devices. Therefore, researchers have challenged themselves to use DNA or RNA to build logical devices for molecular computation. As early as 1994, Leonard Adleman used DNA to solve a computational problem. He encoded the Hamiltonian path problem into different ssDNA sequences and applied biotechnologies (such as Ligation, PCR, Sequencing) to decode the answers to the correct Hamiltonian path.<sup>5</sup>

However, subsequent progress in constructing molecular computing devices has been frustratingly slow. Scientists have demonstrated that DNA-based biocircuitry can perform logic gate operations, signal restoration, amplification, feedback, and cascading, all by distinct DNA strands.<sup>46</sup> But the chief objective is still to emulate or mimic the digital logic found in typical circuit boards with computing power far from rivaling that of silicon computers on the execution of any algorithm. It has become increasingly clear that the ability to interact with naturally occurring biomolecules, together with such unique properties as programmability, nanometric size, and autonomous operation, can be used for practical applications of DNA-based logical devices. This has allowed biological properties to be interfaced to other materials and has opened a novel and exciting direction in biological and biomedical applications.

### **Construction of DNA Circuit: Toehold-Mediated Strand Displacement**

A detailed understanding of nucleic acids, including the specific connectivity of the nucleotides, the Watson-Crick base pairing, and the double-helical structure of the double-stranded DNA (dsDNA), provides the theoretical basis for constructing DNA-based logical devices. To engineer complex logical modules from molecules, the stereotypical structures of these modules should include three components corresponding to input, processor and output. In addition, specific mechanism should be employed to allow the signal transfer through these elements. For DNA-based devices, several common strategies to make circuit construction a predictive science are described.

One significant mechanism to construct a DNA-based logical network is the toehold-mediated displacement reaction,<sup>47</sup> in which a ssDNA (input) “reacts” with a dsDNA to displace one strand (output) of the dsDNA through binding to an overhanging

ssDNA region (toehold) of the dsDNA. As shown in Fig. 1-8A, this reaction involves a three-step process, including toehold binding, branch migration and strand dissociation. Toehold binding occurs via the complementary sequences of the input ssDNA and the toehold. The displacement rate can be adjusted by changing the toehold base numbers. Branch migration is a reversible process with the same binding sequences between the input and output, except that, the remainder of each sequence can be freely designed, allowing the design diversity of a logical network.

Although signal cascading by this classic toehold displacement reaction allows programmability of circuitry design, this methodology is feasible only for simple circuits which allow initial information-processing. One important reason is that the displacement reaction requires longer strand to displace shorter one, resulting decreasing sequence freedom for the downstream strands, thus leading to small amount of circuit layer assembly. In order to improve circuit complexity, people have developed toehold-exchange displacement. (Fig 1-8B) The incoming single strand can be designed with a “shifted sequence” that is complementary to the bottom strand. This can still allow the displacement reaction occur, but the displaced strand will also have a single strand-toehold region that may initiate the backward displacement reaction.<sup>48</sup> This will result in a reversible DNA displacement reaction with the input and output signals of the same sequence length. To drive the reaction in the forward direction, a smart toehold exchange reaction has been developed, in which a fuel strand with much higher concentration creates a catalytic cycle and allows the input sequence to be converted to output without being consumed.<sup>49</sup> The reaction described above is driven forward by the entropy of equilibration for entire circuits. With this mechanism at hand,

Qian et al. constructed DNA-based logical circuits with much highly complexity, including a neural network having 110 DNA strands that demonstrates associative memory capable of answering 81 possible questions,<sup>50</sup> as well as a 130-stranded DNA computer capable of computing the square root of a 4 digit binary number.<sup>51</sup>

### **DNA-based Circuit for Analytical and Biomedical Applications**

Signal amplification is an important strategy in the development of molecular systems with increased responsiveness. Recent advances in the field of nucleic acids have generated DNA-based circuits, in which enzyme-free signal amplification can be achieved by simple nucleic acid hybridization. Compared to traditional enzyme-based amplification, pure DNA-based amplification has the advantages of robust adaptability, flexible engineering and the feasibility for intracellular applications. Dirks and Pierce demonstrated an isothermal, enzyme-free method for highly sensitive detection of a particular DNA strand.<sup>52</sup> The protocol made use of multiple copies of two distinct DNA hairpins that are initially stable without any cross-hybridization. A ssDNA sequence is able to initiate hybridization between the two hairpins, allowing the cascade effect to occur completely and autonomously. In addition, Zhang et al. developed a general methodology for designing amplification circuits with pure DNA molecules by using entropy-driven catalytic reactions.<sup>53</sup> Such DNA logical circuits have applications to a variety of analytical procedures, in which a larger response (e.g., a cascade response) is needed as output in response to one of multiple molecular detection events.

In addition to their analytical applications, researchers increasingly envision an important role for artificial DNA-based circuits in biological engineering. For example, DNA-based circuits could be utilized as a type of servomechanism to manipulate the functions of biological molecules *in vitro*<sup>54</sup> and regulate gene expressions *in vivo*.<sup>55</sup> Also,

as discussed above, because of their roles in monitoring biological reactions, functional nucleic acids provide excellent platforms to bridge nucleic acid-based circuits to other biological components (such as proteins and carbohydrates). By involving functional nucleic acid components into a circuit, smart and flexible manipulation of their functions can be realized, thus extending the biological and biomedical applications of nucleic acid-based circuitry. (Fig.1-9)

### **DNA-based Biomimicry**

The design of artificial chemical alternatives that mimic living systems, especially in regard to complex self-organizing systems, has been an effective means to elucidate biological process. Likewise, natural systems also provide us with insight into the design and construction of artificial mimicry systems. For instance, inspired by the self-assembly process in nature, the strategy of “bottom-up” was intensively applied in the construction of various nanostructures using molecular interactions without external guidance. However, to mimic complex natural system, many challenges need to be overcome, including controlling programmability and systemization, as well as transferring high-level biological codes into simple artificial modules and physical implementations. Another difficulty involves forming robust molecular structures and precisely controlling their temporal dynamics. As discussed above, building on the richness of DNA logical circuit construction, synthetic DNA-based systems have been used to explore the possibilities of mimicking both simple and complex systems. For example, small-scale *in vitro* circuits encoding elementary functions, such as counters,<sup>56</sup> bistable memory,<sup>57</sup> or oscillations,<sup>58</sup> have successfully been engineered. In addition, larger networks, including mimicry of the neural network<sup>50</sup> and predator-prey ecosystems<sup>59, 60</sup> continue to demonstrate the excellence of DNA biochemistry for

biomimicry implementation. These successful investigations definitely will broaden the scope of using DNA to mimic complex natural systems for various applications.

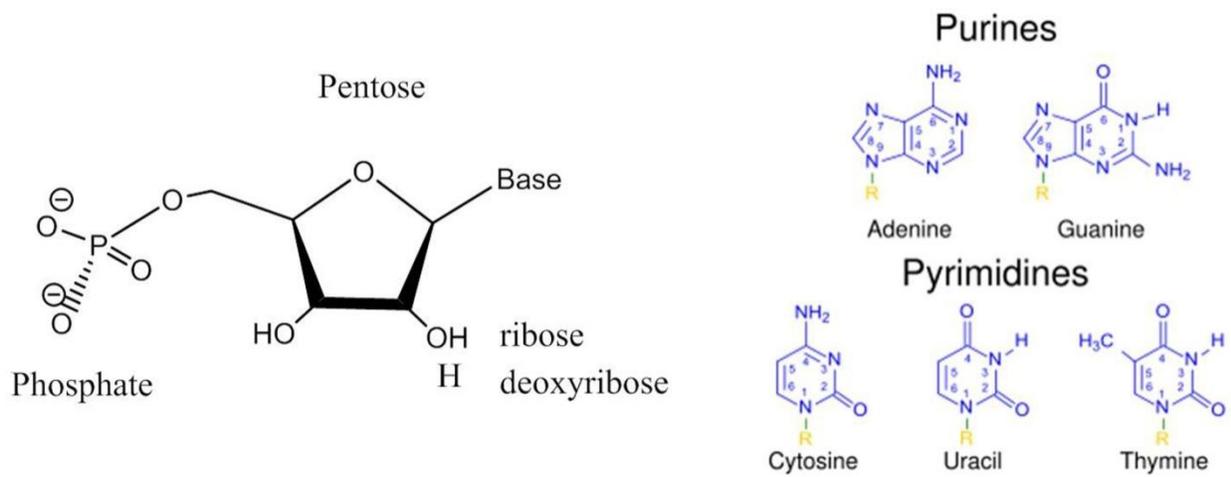
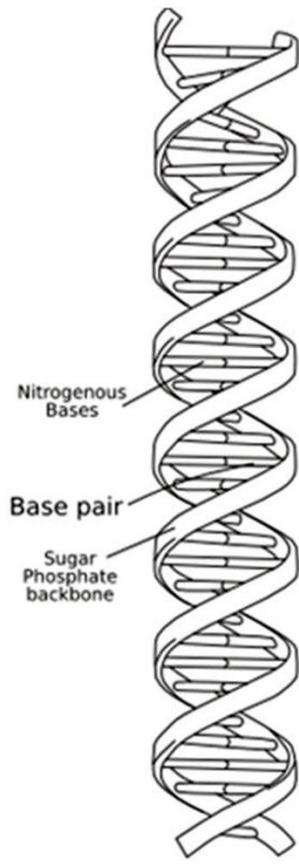


Figure 1-1. Chemical structure of nucleic acids and nucleosides.



**DNA**

Deoxyribonucleic acid

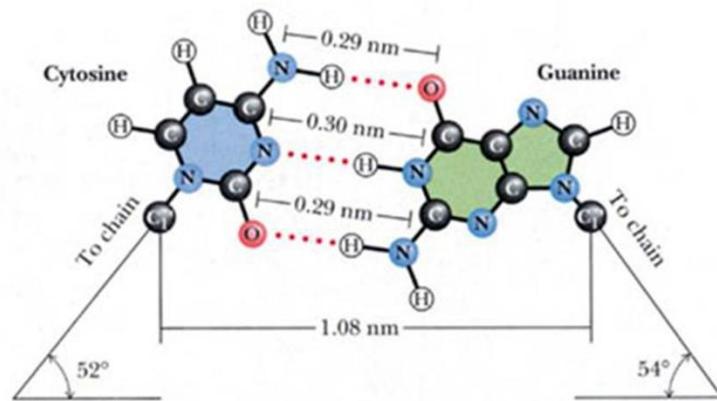
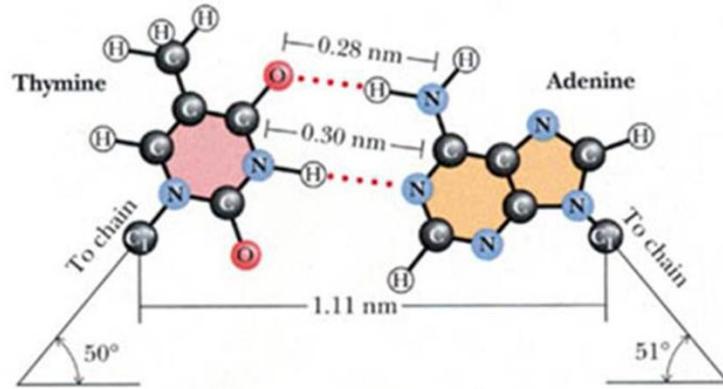


Figure 1-2. Structure of DNA duplex and scheme of Chargaff's rule.

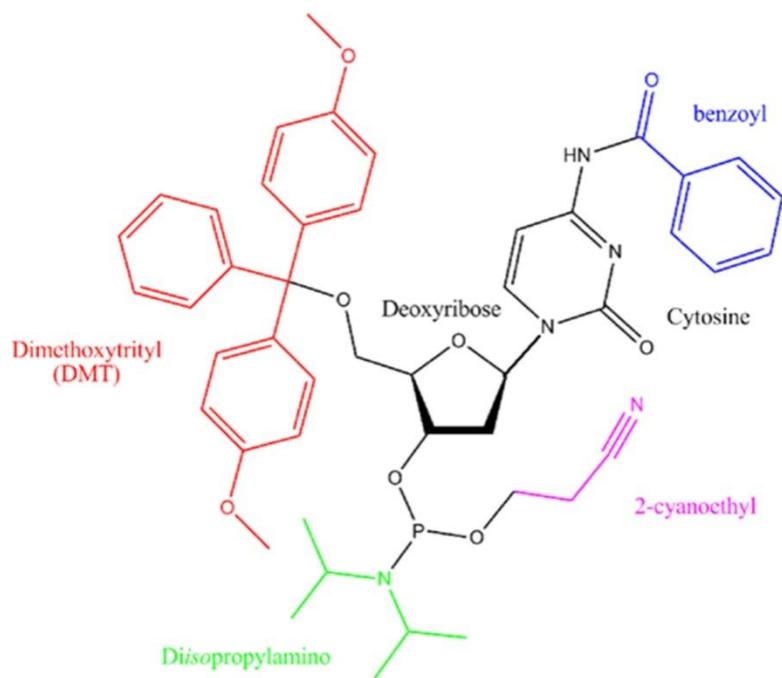


Figure 1-3. Chemical structure of DNA phosphoramidite.

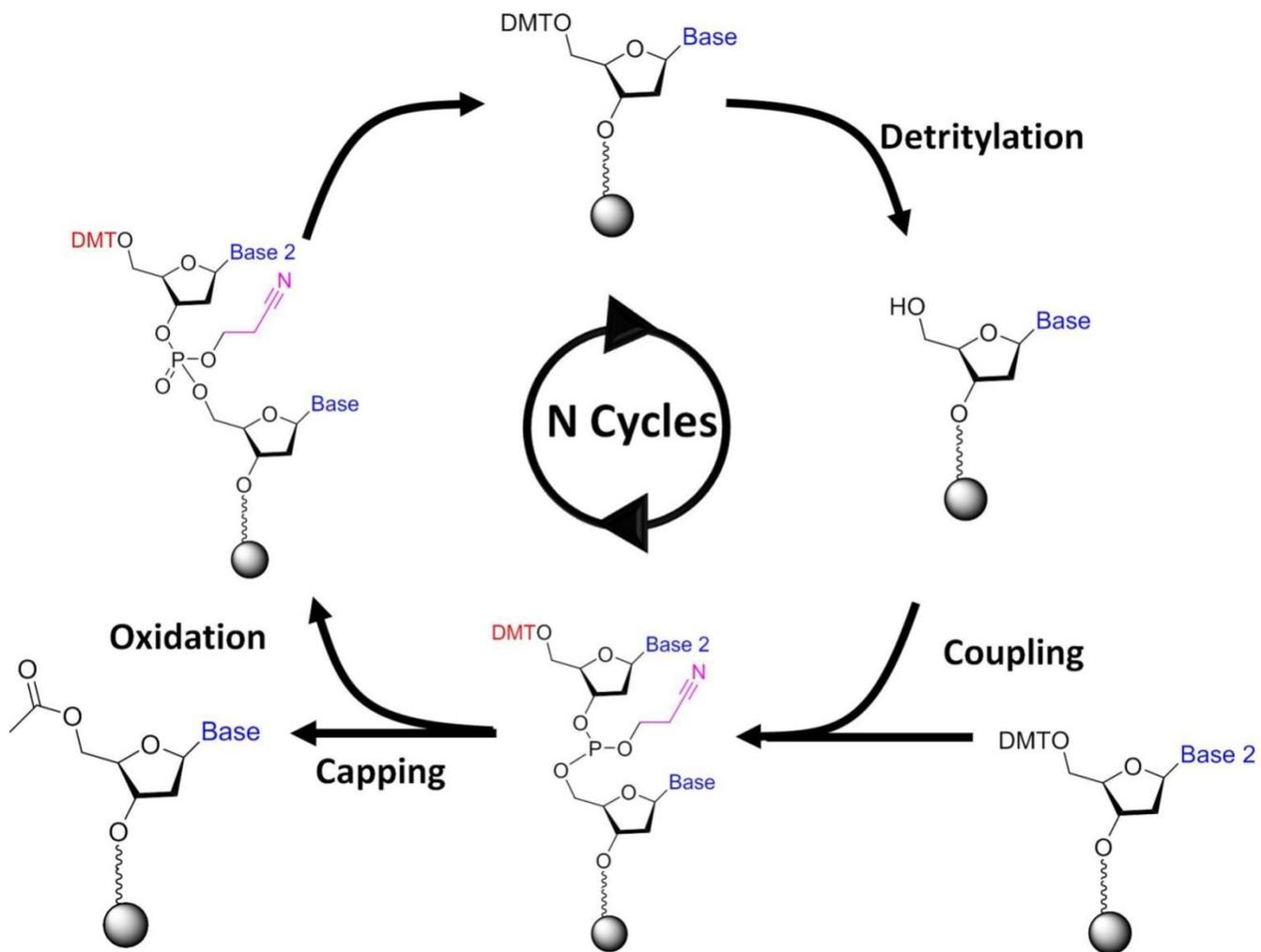


Figure 1-4. Automated oligonucleotide synthesis based on phosphoramidite chemistry.

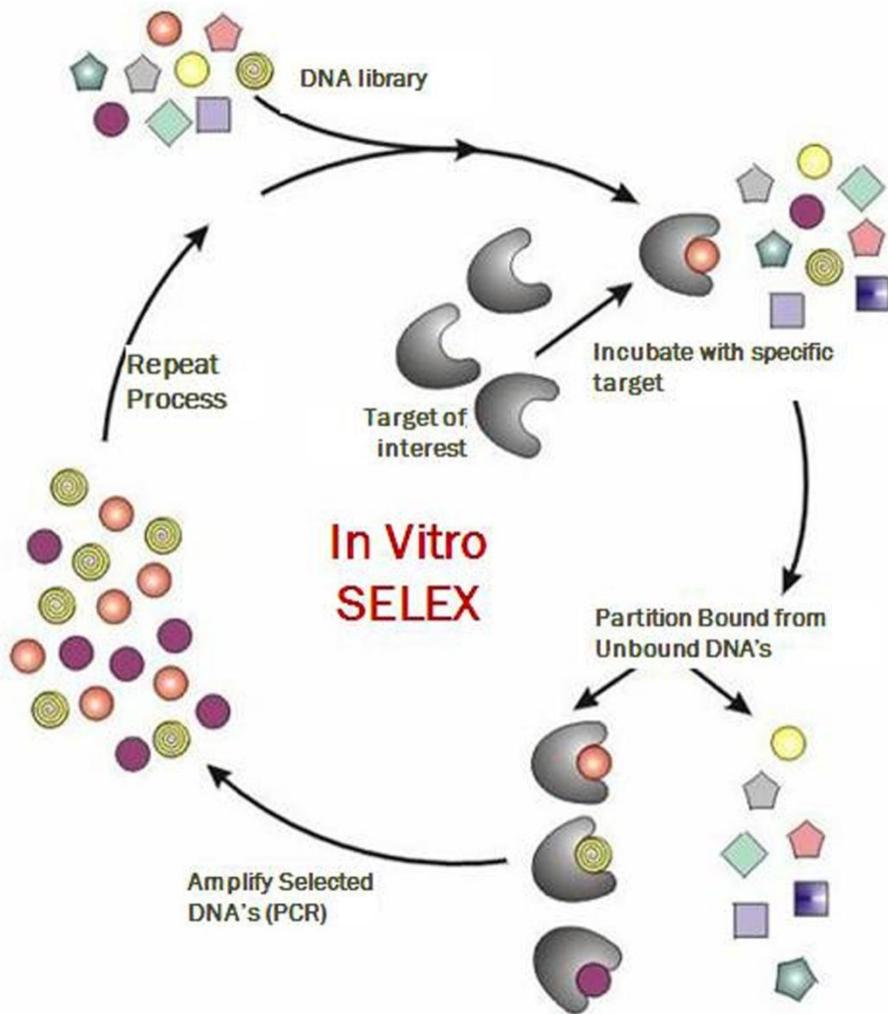


Figure 1-5. Scheme of Systematic Evolution of Ligands by EXponential enrichment (SELEX).

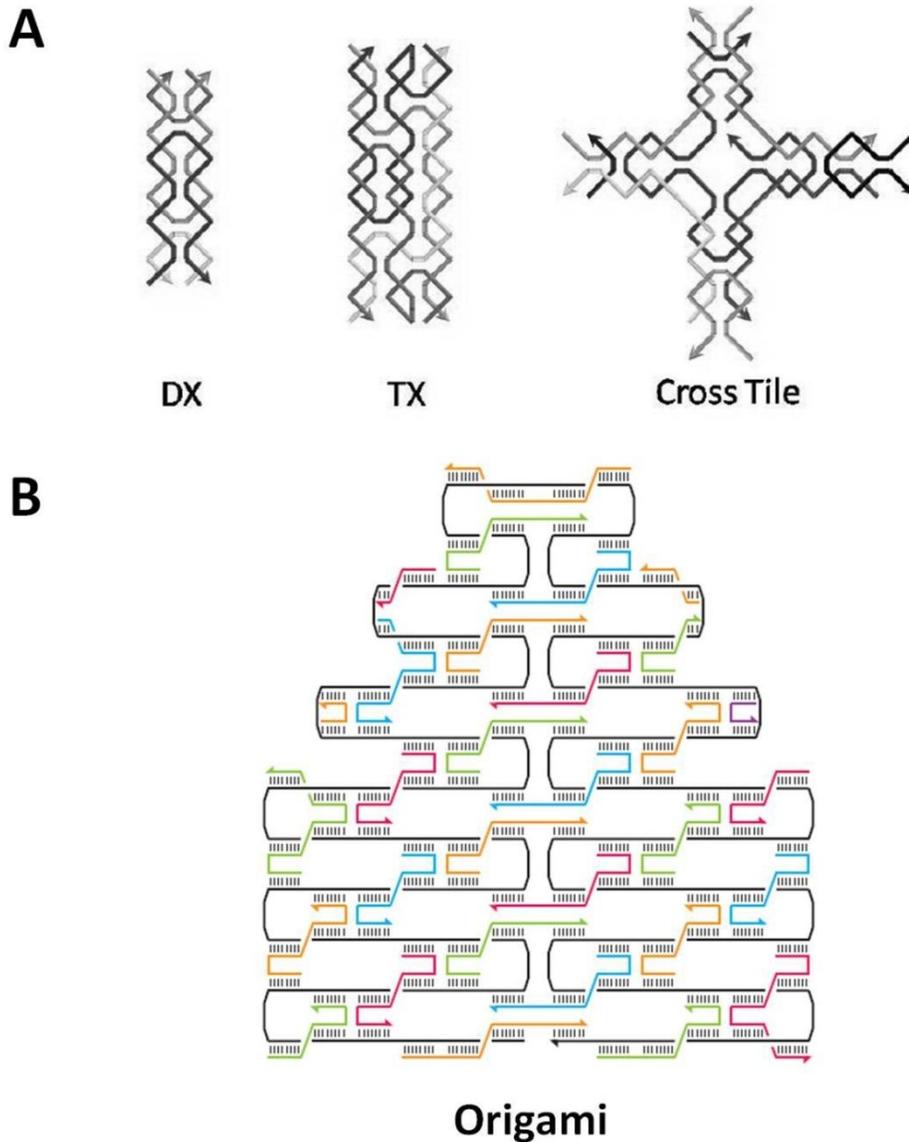


Figure 1-6. A) Illustration of some DNA tiles.<sup>61</sup> B) Schematic folding of DNA origami. Reprinted by permission from Paul Rothemund, Folding DNA to create nanoscale shapes and patterns (Page 298, Figure 1-e), Nature 440, Page 297-302, 2006, Macmillan Publishers Ltd, New York, USA.

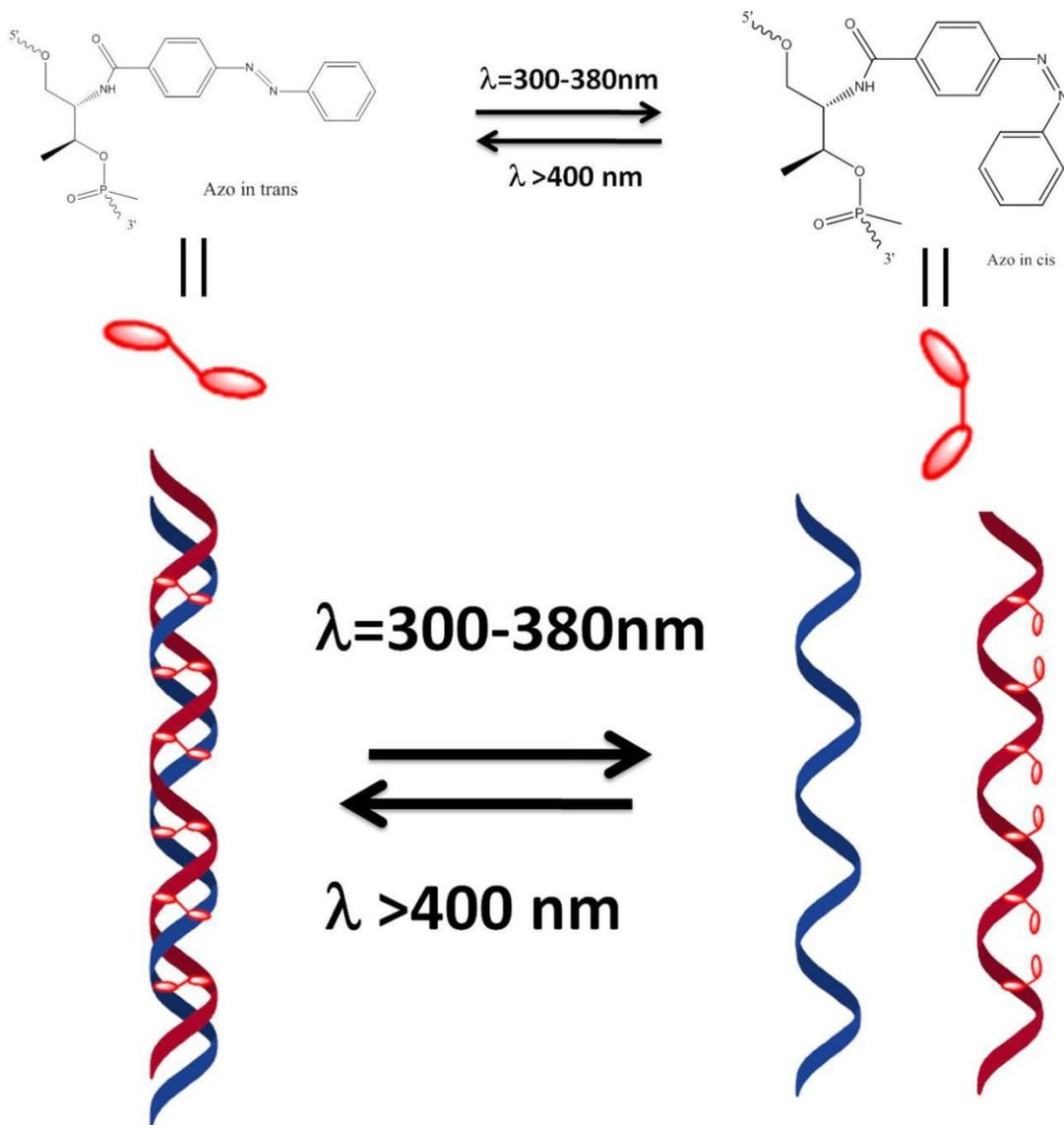


Figure 1-7. Scheme of the reversible hybridization/dehybridization of an azobenzene incorporated DNA duplex.

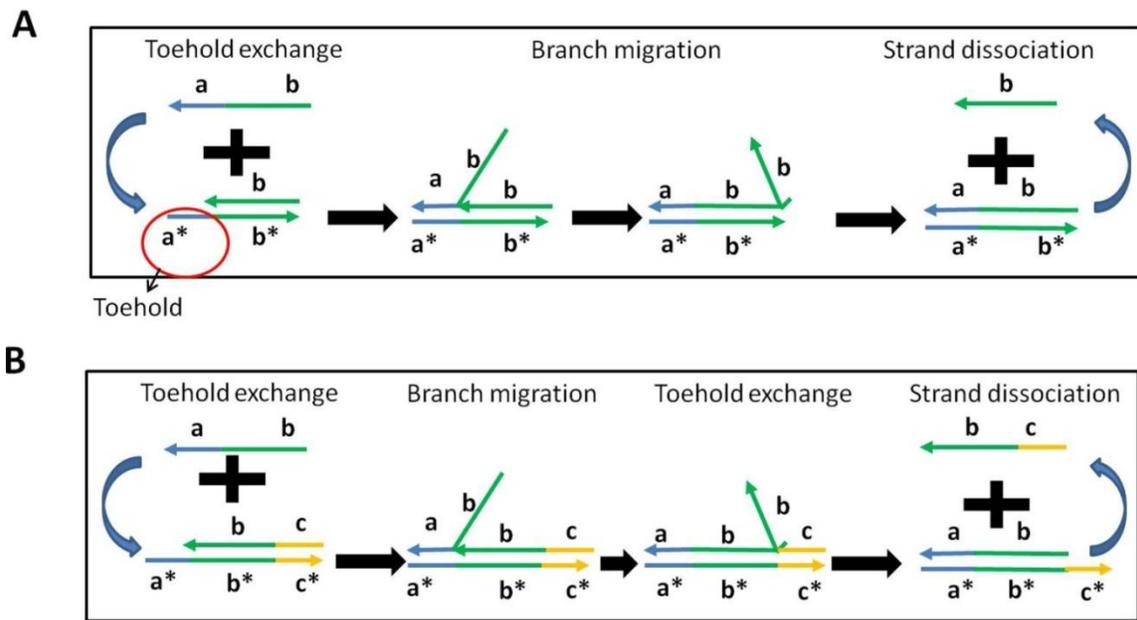


Figure 1-8. A) Scheme of toehold-mediated displacement reaction. B) Scheme of toehold-exchange displacement reaction.

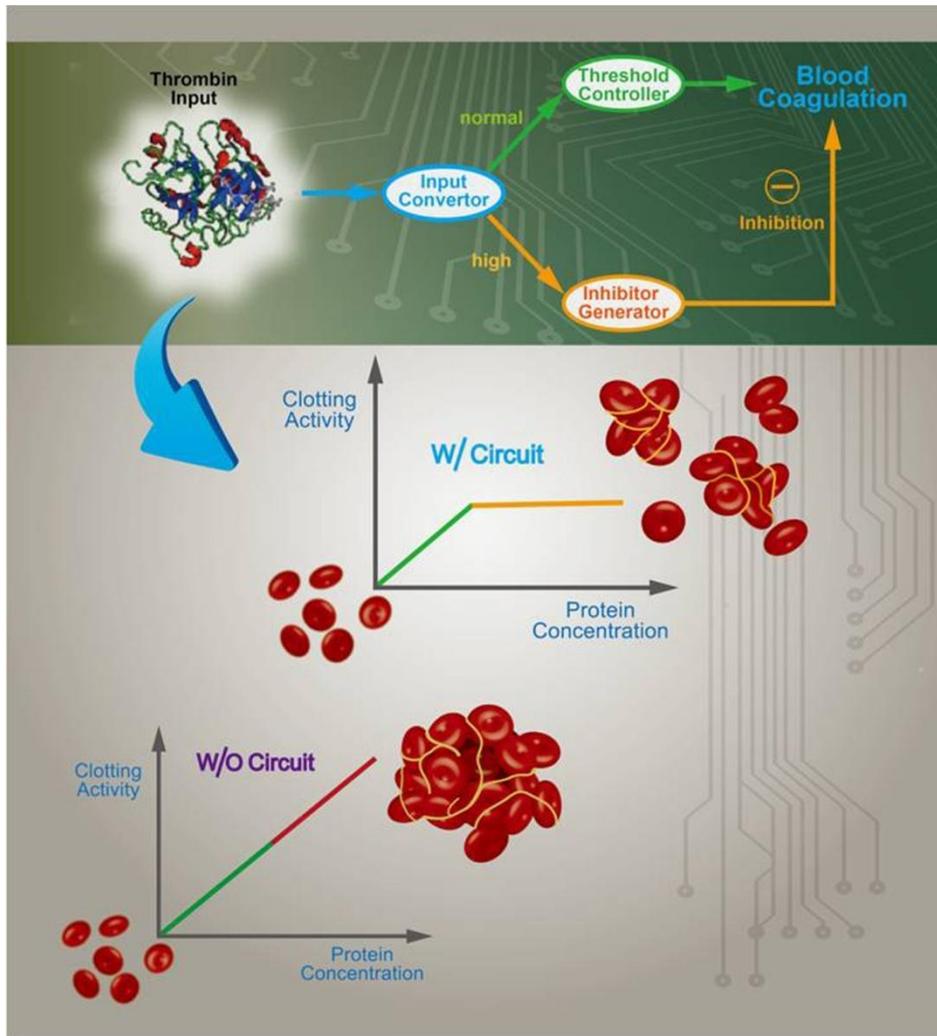


Figure 1-9. Example of functional nucleic acid-based circuit for biological and biomedical applications

## CHAPTER 2 MOLECULAR ENGINEERING OF A CELL-SURFACE APTAMER CIRCUIT FOR TARGETED AND AMPLIFIED PHOTODYNAMIC CANCER THERAPY

### **Significance and Background**

As a minimally invasive therapeutic modality, photodynamic therapy (PDT) is already greatly used in clinical treatment of cancers. PDT can destroy cancer cells when light irradiates a photosensitizer (PS), generating reactive singlet oxygen ( $^1\text{O}_2$ ).<sup>62</sup> Briefly, PDT involves a two-step process, whereby a nontoxic PS is delivered to an organism and then activated by an appropriate light source. However, because the  $^1\text{O}_2$  has a limited lifetime and diffusion distance, efficient and reliable PDT depends on generating  $^1\text{O}_2$  with methods that offer the greatest selectivity.<sup>63</sup>

PDT selectivity is usually controlled at two levels. The first level controls the spatial localization of PS reagents. This approach has been actively pursued by specifically delivering PS to the tumor site with regional light shining, which has effectively improved PDT selectivity and efficiency.<sup>64, 65</sup> But the tendency to cause damage to surrounding normal tissues still exists. To achieve greater selectivity, a molecular activation layer is added to further control the specificity of the PS. At this level, the probe initially stays in the nontoxic state and can only be activated when it interacts with its corresponding trigger at the tumor site. For example, we and others have developed activatable PDT methods which can be triggered by biomarkers, including membrane proteins<sup>66</sup> and extracellular proteases,<sup>67</sup> as well as cellular environments (e.g., pH)<sup>68</sup> or other external stimuli, including artificial molecular switches.<sup>69-71</sup> Molecular activation allows the PS to distinguish diseased from healthy cells, thus greatly improving the selectivity of PDT.

To achieve higher oncolytic efficacy in tumors, a sufficient dosage of drugs should be administered at the tumor site. However, greater selectivity is typically achieved by introducing multiple activation processes, but at the cost of decreasing the active drug amounts at the tumor site. For example, the triggers of activatable PDT are usually biomarkers present inside or outside the cells. However, the limited amount of trigger elements in the disease cells coupled with low activation efficiency may dramatically decrease the activation and killing effects of the PS. Selective amplification would effectively solve this problem. Researchers have applied some enzymes, such as protease whose overexpression is correlated with specific diseases, to continually catalyze PS activation, thereby amplifying the PDT effects.<sup>72</sup> However, the application of enzymes is often limited by their microenvironment, including pH and temperature, in turn reducing the applications of enzyme-activatable PDT. Thus, to achieve more robust PDT with selectivity and amplification effect, a suitable medium is required that can both recognize the target cell and amplify the therapeutic effect.

As carriers of genetic information with well-regulated and predictable structures, nucleic acids are promising materials for signal amplification based on their nanometer size and programmability. Recent advances in the field of nucleic acids have generated nucleic-acid based circuits, in which enzyme-free signal amplification can be achieved by simple nucleic acid hybridization, such as hybridization chain reaction,<sup>52, 73</sup> entropy-triggered hybridization catalysis<sup>49</sup> and DNA hairpin fuel catalysis.<sup>74, 75</sup> These methods show promise in amplifying PDT with such properties as high amplification efficiency, environmental robustness and ability to interact with other naturally occurring molecules. Meanwhile, the exploration and development of special single-stranded

oligonucleotides, well known as aptamers,<sup>10, 76, 77</sup> have extended the recognition capabilities of nucleic acids from Watson-Crick base-pairing to interactions with various targets, such as small molecules,<sup>78</sup> proteins,<sup>79</sup> and cells,<sup>80</sup> via the aptamers' unique secondary or tertiary structures. We recently developed an effective method to generate aptamer-based molecular probes for the specific recognition and targeting of cancer cells.<sup>81-83</sup> Therefore, by combining the recognition and amplification abilities of these oligonucleotides, more efficient and specific PDT methods can be developed.

In this chapter, we report the design of an aptamer-based DNA circuit capable of the selective recognition of cancer cells, controllable activation of PS and amplification of therapeutic effect. In particular, the amplification circuit motif comprises two DNA hairpins developed by Yin et al.<sup>74</sup> and Li et al.<sup>75</sup> In principle, two DNA hairpin structures,  $A_1$  and  $A_2$  initially do not hybridize with each other because of the effective block created by complementary domains. However, in the presence of another ssDNA sequence, termed catalyst (C),  $A_1$  and  $A_2$  can form a stable duplex without consuming C. As shown in Figure 2-1,  $A_1$ ,  $A_2$  and C contain a few functional domains labeled in lowercase letters. Complementarity between lettered domains is denoted by an asterisk. Initially, C can hybridize with the exposed toehold domain a of  $A_1$  and gradually open the stem of  $A_1$  to form intermediate  $A_1C$ , but  $A_1C$  has an exposed ssDNA domain  $c^*$  able to hybridize with the exposed domain c in  $A_2$ . Hence, after hybridization of c and  $c^*$ , the sequence  $dc^*b^*d$  will undergo branch migration and displace the C sequence ( $c^*b^*a^*$ ) to form the  $A_{12}$  duplex. Importantly, the released C triggers further hybridizations of  $A_1$  and  $A_2$  in repeating cycles, thus providing the multiple-trigger effect absent in previous models. In this example, C catalyzes the formation of duplex  $A_{12}$

from  $A_1$  and  $A_2$  through a prescribed reaction pathway. The overall reaction is driven by a decrease in enthalpy resulting from the formation of  $A_{12}$  with a greater number of base pairs.

Selectivity is achieved by encoding the catalyst sequence C ( $a^*b^*c^*$ ) into an aptamer sequence (Apt-C) that targets cancer cells. To avoid forming undesired secondary structures, 17 poly T bases are used to separate the aptamer sequence and C. Under these conditions, the aptamer part can bind to the receptor on the target cancer cell membrane with a tail (C) exposed for the catalytic hybridizations of  $A_1$  and  $A_2$ . Thus, one aptamer binding event can induce multiple hybridization events between  $A_1$  and  $A_2$  to form dsDNA  $A_{12}$ , and the all-important amplification step essentially derives from the catalytic reaction.

To apply this cell-catalyzed hairpin amplification circuit to PDT therapy, dsDNA sequences denoted as  $R_{12}$  are employed to carry the photodynamic therapeutic reagents. Because of its high photosensitizing efficacy and low dark toxicity, Chlorin e6 (Ce6), a second-generation and easily modifiable photosensitizer, is modified on ssDNA  $R_1$ , and a quencher, BHQ2, is conjugated on ssDNA  $R_2$  to quench the generation of  $^1O_2$  by Ce6 when no target cell is present. This design has several advantages. First, specific binding-induced activation allows the DNA circuit to distinguish diseased cells from healthy cells, reducing damage to nearby healthy cells which otherwise might be destroyed during PDT with conventional photosensitizers. Second, since the circulatory system *in vivo* can flush away the unbound aptamers and the catalyst C with aptamer is only present on the target cell membrane, the catalytic reaction will only take place close to the target cancer cells, resulting in a high local concentration of  $^1O_2$  to

selectively kill the target cells. Third, traditional aptamer-based PDT has suffered from the drawback of insufficient killing effects from the 1:1 binding-induced singlet oxygen generation (SOG). While, in this new design, by incorporating the catalyst sequences of the hairpin amplification circuit on cell membranes, numerous binding-induced SOG events can be realized on each cell membrane. In addition, the uncatalyzed background of the circuit is nearly undetectable, resulting in fewer side effects to other healthy cells. Finally, to our best knowledge, this is the first design using the target cancer cell as the trigger to drive the DNA hybridizations. As such, this method may provide a universal strategy for signal amplification on cell membranes.

## **Experimental Materials and Methods**

### **Cell Culture**

Ramos (CRL-1596, B-cell line, human Burkitt's lymphoma) and CCRF-CEM (CCL-119, T-cell line, human acute lymphoblastic leukemia) were cultured in RPMI 1640 medium (American Type Culture Collection) with 10% fetal bovine serum (FBS, Invitrogen, Carlsbad, CA, USA) and 0.5 mg/mL penicillin-streptomycin (American Type Culture Collection) at 37°C under a 5% CO<sub>2</sub> atmosphere. Cells were washed before and after incubation with washing buffer [4.5 g/L glucose and 5 mM MgCl<sub>2</sub> in Dulbecco's PBS with calcium chloride and magnesium chloride (Sigma-Aldrich)]. Binding buffer used for selection was prepared by adding yeast tRNA (0.1 mg/mL; Sigma-Aldrich) and BSA (1 mg/mL; Fisher Scientific) to the wash buffer to reduce background binding.

### **Ce6-Modified DNA Synthesis**

An amino group (Glen Research Corp.) was incorporated at the 5'-end using the synthesis protocol specified by the company. After removing the MMT protection group on the 5'-amino of the sequence on machine, the CPG beads were washed with

acetonitrile (ACN) 10 times and dried with nitrogen for off-machine coupling of Ce6. Each Ce6 molecule has three carboxyl groups for conjugation with the amino group. To improve the coupling efficiency and reduce the multiple coupling products, the amount of Ce6 was 10 times more than DNA product in the coupling reaction. Ten  $\mu\text{mole}$  Ce6 was mixed with an equal molecular amount of N,N'-dicyclohexylcarbodiimide (DCC, Sigma-Aldrich, Inc.) and N-hydroxysuccinimide (NHS, Sigma-Aldrich, Inc.) and dissolved in 250  $\mu\text{L}$  N,N-Dimethylformamide (DMF) for the activation reaction with 1 hr stirring. The coupling reaction was performed with vigorous stirring overnight, followed by washing with ACN. Then the DNA product was purified by HPLC.

### **DNA Purification**

Native PAGE was applied to purify the  $A_1$ ,  $A_2$  hairpin strands to remove excess strands and avoid undesired system leakage.  $A_1$  and  $A_2$  were annealed at concentrations of around 50  $\mu\text{M}$  in 1 $\times$  TAE-Mg buffer (40 mM Tris-Acetate-EDTA, pH 8.0, 12.5 mM  $\text{Mg}(\text{Ac})_2$ ) and cooled to room temperature. Native PAGE gels (12%) in 1 $\times$  TAE-Mg buffer were run at 110 V for 90 minutes at 4 $^\circ\text{C}$  and stained with GelRed stain solution (Biotium, Inc., Hayward, CA). Only the sharp bands were cut from the gels, chopped into small pieces, and soaked in 1 $\times$  TAE- $\text{Mg}^{2+}$  buffer for 24 hr. The buffer was extracted and concentrated with centrifugal filter devices (Millipore, Billerica, MA). Finally, the purified DNA sequences were quantified by UV spectrometry and kept in buffer for future use.

### **Fluorescence Kinetics of DNA Hairpin Circuit in Buffer**

All fluorescence measurements were performed using a Fluorolog (Jobin Yvon Horiba) with a 100  $\mu\text{L}$  macro cuvette. DNA sequences  $A_1$ ,  $A_2$  were separately refolded

in the Fluo buffer (20mM Tris, pH 7.5; 140mM NaCl; 5mM KCl). This and other refolding reactions involved heating to 90°C for 1 min, followed by slowly decreasing the temperature to 25°C at a rate of 0.1°C s<sup>-1</sup>. After purification by gel electrophoresis, the annealed sequences were stocked in Fluo buffer for later use. An amount of 10 μM stock of R was prepared by annealing 10 μM FAM-R<sub>1</sub> and 15 μM DABCYL-R<sub>2</sub> in Fluo buffer. An excess of R<sub>2</sub> ensures efficient quenching of R<sub>1</sub>, but does not interfere with the readout of A<sub>1</sub> and A<sub>2</sub>. A mixture of 100 nM A<sub>1</sub>, 100 nM A<sub>2</sub> and 150 nM R<sub>12</sub> was prepared in 1×Fluo buffer. The fluorescence at 518 nm was monitored at 25 °C after adding different amounts of TDO5-C. To evaluate the amplification effect of the circuit, a 1:1 displacement reaction was performed under the same conditions by mixing 10 μM A<sub>1</sub> and A<sub>2</sub> in 1×Fluo buffer and heating to 90°C for 3 minutes, followed by slowly decreasing the temperature to 25 °C to form the stable duplex A<sub>12</sub>. A 150 nM sample of R<sub>12</sub> was incubated in 1×Fluo buffer, followed by adding different concentrations of A<sub>12</sub> and monitoring the fluorescence.

### **Test of Fluorescence Response and SOG of DNA Hairpin Circuit**

For these experiments, Ce6-modified R<sub>1</sub> and BHQ2-modified R<sub>2</sub> were used to form duplex R<sub>12</sub>. To study the Ce6 fluorescence response to different concentrations of TDO5-C, 100 nM A<sub>1</sub>, 100 nM A<sub>2</sub> and 150 nM R<sub>12</sub> were mixed in 1×Fluo buffer. The excitation wavelength was set at 404 nm with emission scanned from 600 nm to 800 nm. When testing the SOG, the concentration of SOSG probe was set at 2 μM together with 100 nM A<sub>1</sub>, 100 nM A<sub>2</sub> and 150 nM R<sub>12</sub> in 1×Fluo buffer. To extend the lifetime of <sup>1</sup>O<sub>2</sub> and increase the sensitivity of SOG assay, all buffers and samples were prepared using deuterium oxide. The SOG was triggered by irradiation at 404 nm, the maximum

absorption of Ce6, for 10 minutes. The SOSG fluorescence was obtained with excitation at 494 nm and emission from 500 nm to 600 nm.

### **Test of SOG Response of DNA Hairpin Circuit to Cancer Cells**

$3 \times 10^5$  of Ramos and CCRF-CEM cells were prepared in 100  $\mu\text{L}$  of washing buffer separately. Fifty pmol of nonlabeled TDO5-C were added and incubated for 30 minutes. After washing the cells twice with washing buffer, the two different cell types were resuspended in 100  $\mu\text{L}$  of washing buffer. Then 100 nM  $A_1$ , 100 nM  $A_2$  and 150 nM  $R_{12}$  were incubated with cells for 1 hr. To extend the lifetime of  $^1\text{O}_2$  and increase the sensitivity of SOG assay, all buffers and samples were prepared using deuterium oxide. Two micromolar SOSG sensors were added to the cell medium, and SOG was triggered by irradiation at 404 nm, the maximum absorption of Ce6, for 10 minutes. The fluorescence was monitored at 25 °C with the excitation wavelength at 494 nm and emission from 500 nm to 600 nm.

### **Flow Cytometric Analysis**

In flow cytometry tubes, 250 nM Biotin-labeled TDO5 or TDO5-C was incubated with  $3 \times 10^5$  Ramos or CCRF-CEM cells at 4°C in 200  $\mu\text{L}$  of binding buffer for 30 minutes. The cells were washed twice with 1 mL of washing buffer, centrifuged at 1300 rpm for 3 min, and then resuspended in 200  $\mu\text{L}$  of washing buffer. One  $\mu\text{L}$  (1:400) Streptavidin-conjugated PE dye was incubated with the cells for another 20 minutes and washed twice using washing buffer. The cells were analyzed on a FACScan Flow Cytometer by counting 30,000 events. The PE-labeled unselected ssDNA library was used as a negative control.

## **Confocal Imaging of Cells Bound with Aptamer**

For confocal imaging, the Ramos and CCEF-CEM cells were incubated with 50 pmol of TAMRA-labeled TDO5-C in 100  $\mu$ L of binding buffer containing 20% FBS on ice for 30 min. The cells were washed twice with 1 mL of washing buffer, centrifuged at 1300 rpm for 3 min, and then resuspended in 100  $\mu$ L of washing buffer. Twenty microliters of cell suspension bound with TAMRA-labeled TDO5-C were dropped on a thin glass slide placed above a 60 $\times$  objective on the confocal microscope. Imaging of the cells was performed on an Olympus FV500-IX81 confocal microscope. A 5-mW, 543-nm He-Ne laser was the excitation source for TAMRA throughout the experiments. The objective used for imaging was a PLAPO60XO3PH 60 $\times$  oil-immersion objective with a numerical aperture of 1.40 (Olympus).

## **Cytotoxicity Study**

The cytotoxicity study was performed using the CellTiter 96 Aqueous One Solution cell proliferation assay (MTS) for Ramos and CCRF-CEM cell lines in a 96-well cell culture plate at 500k/well, 100  $\mu$ L. Five groups of cell samples were set up as follows: group 1, cells only; group 2, cells+ 3.3  $\mu$ L 30  $\mu$ M A<sub>1</sub>, 1.3  $\mu$ L 77  $\mu$ M A<sub>2</sub> and 2  $\mu$ L 50  $\mu$ M Ce6-modified R<sub>12</sub>; group 3, cells incubated with the same amount of A<sub>1</sub>, A<sub>2</sub> and R<sub>12</sub>, together with 1  $\mu$ L 10  $\mu$ M free C; group 4, cells incubated with 1  $\mu$ L 100  $\mu$ M preannealed A<sub>12</sub> and 2  $\mu$ L 50  $\mu$ M Ce6-modified R<sub>12</sub>. For group 5, the cells were incubated with TDO5-C conjugates for 30 min at 4 $^{\circ}$ C, followed by centrifugation at 1300 rpm for 3 min to remove the unbound DNA. Then probes in the same amount as that of group 2 were added. All groups of cells were suspended in cell medium (No FBS) and then irradiated with white light on ice for 3 hr. After irradiation, the cells were incubated in a CO<sub>2</sub> incubator for 36 hr. Finally, a 6 $\times$ -concentrated MTS solution (120  $\mu$ L/well) in

RPMI 1640 medium solution was added to each well and incubated at 37 °C for 2 h. The absorbance value at 490 nm was determined by a VersaMax microplate reader (Molecular Devices, Inc., Sunnyvale, CA).

## Results and Discussion

### Signal Amplification Effect of Aptamer Circuit in Buffer

To demonstrate the effectiveness of the C ( $a*b*c^*$ ) in catalyzing the  $A_1$  and  $A_2$  hybridization, native gel electrophoresis was used. As shown in Figure 2-2, without C,  $A_1$  and  $A_2$  can be present stably without hybridization. However, when C is added,  $A_1$  and  $A_2$  hybridize with each other to form  $A_{12}$  with a yield even higher than that achieved by annealing of  $A_1$  and  $A_2$ . To further study the amplification efficiency of the hairpin circuit, a FRET-based dsDNA  $R_{12}$  was designed with a stable fluorophore (FAM- $R_1$ ) and quencher (DABCYL- $R_2$ ) pair. To improve the thermostability and anti-enzymatic digestion ability of the  $R_{12}$  duplex, we incorporated 4 LNA (Locked nucleic acid) nucleotides into FAM-labeled  $R_1$ .<sup>84</sup> C was linked to an aptamer sequence TDO5,<sup>81</sup> which targets acute lymphoblastic leukemia B-cells ( $K_d=74.7$  nM) via a poly-T linker. Thus, TDO5-C was used as catalyst to initiate the  $A_1/A_2$  hybridization, and the fluorescence was monitored. In the presence of different concentrations of TDO5-C (0-20 nM), dramatic signal enhancement was observed (Fig. 2-3A), indicating the effective catalytic effect of TDO5-C. In addition, the signals approached a maximum in 2 hr, indicating rapid kinetics of the catalytic hybridization.

We also studied the fluorescence kinetics of the 1:1 displacement reaction (Fig. 2-3B). In particular,  $A_{12}$  was prepared by annealing equal concentrations of  $A_1$  and  $A_2$  in advance. Then, different concentrations of  $A_{12}$  were added to displace  $R_{12}$  in buffer

solution. In Figure 2-3C, the fluorescence enhancement ratio  $\left(\frac{F_{1:n} - B}{F_{1:1} - B}\right)$  between catalytic amplification circuit (1:n) and displacement reaction (1:1) is 8-fold at the target concentration of 20 nM after about 2 hr, indicating high amplification efficiency of this circuit.

For therapeutic applications with the DNA circuit,  $A_1$ ,  $A_2$  and  $R_{12}$  will be present together in the deactivated forms around cells, and small leakage hybridizations may occur. Therefore, we tested the leakage hybridization rate by measuring the fluorescence of buffer solution containing  $A_1$ ,  $A_2$  and  $R$  without TDO5-C for 8 hr (Fig. 2-4). Although a small leakage did occur, the result indicated that the second-order rate constant of uncatalyzed reaction could be estimated to be  $<10 \text{ M}^{-1}\text{s}^{-1}$ , which is almost negligible for PDT applications.

Next, the photosensitizer Ce6 was conjugated to the ssDNA  $R_1$ , and the BHQ-2 quencher was modified with ssDNA  $R_2$ . Because of the close proximity between Ce6 and BHQ-2, up to 95% quenching efficiency of Ce6 was observed by our previous studies.<sup>85</sup> Herein, the DNA hairpin circuit had significant fluorescence enhancement upon the addition of different concentrations of TDO5-C. This was illustrated by the Ce6 fluorescence which increased up to 10-fold with 20 nM TDO5-C in buffer (Fig. 2-5). To evaluate the effect of different concentrations of TDO5-C on the amount of  $^1\text{O}_2$  generated by Ce6-modified  $R_1$ , singlet oxygen sensor green (SOSG) was added, and its fluorescence enhancement was measured before and after irradiation at 404 nm. As shown in Figure 2-3D, SOSG fluorescence increased 3-fold with the introduction of 20 nM TDO5-C in the span of 1 hr, indicating that SOG could be mediated by TDO5-C.

## Selective Recognition Ability of the Aptamer Circuit

For proof of concept, a leukemia cell line was chosen as the target. Compared with solid tumor cells, leukemia cells are widespread in the circulatory system and are surrounded by normal blood cells. Under these circumstances, any nonspecific cytotoxin would also destroy the normal blood cells. Therefore, a therapeutic method which can selectively recognize and kill the target leukemia cells is highly desirable. As mentioned above, aptamer TDO5, which binds to the cancer cell membrane protein IgM with high affinity and selectivity, was used in our study with target cancer cell Ramos (acute lymphoblastic leukemia B-cells) and negative control cell CCRF-CEM (acute lymphoblastic leukemia T-cells). Therefore, if TDO5-C is present, it recognizes the target cancer cells and, importantly, also catalyzes the DNA hairpin hybridization to trigger SOG around cells. First, the selective binding of TDO5-C to Ramos cells was demonstrated by flow cytometry, as shown in Figure 2-6 A and B. Herein, aptamer TDO5 was used as positive control. Compared with TDO5, TDO5-C showed almost equally strong binding affinity to Ramos cells at 4°C, indicating that the cell membranes were partially covered by TDO5-C. However, both TDO5 and TDO5-C exhibited weak affinity to the control CCRF-CEM cells, as evidenced by only small fluorescence peak shifts. In addition, to confirm that TDO5-C was bound to the cell membrane surface without internalization, confocal microscopy images were taken with TMR-labeled TDO5-C incubated with Ramos and CCRF-CEM cells (Fig. 2-6 C and D). Since only the cell membrane surface was labeled with fluorescence, the strong binding and low uptake efficiency of TDO5-C makes it suitable for the catalysis of the hairpin circuit in close proximity to the target cancer cells.

### **Singlet Oxygen Amount with Different Cancer Cells**

To determine whether the amplification effect of TDO5-C remains active on the cell membrane, SOG was evaluated by incubating TDO5-C-labeled cells with the circuit ( $A_1$ ,  $A_2$  and  $R_{12}$ ) in PBS buffer, followed by adding SOSG sensor and monitoring the fluorescence. As shown in Figure 2-7, obvious SOSG fluorescence enhancement was observed when incubating the circuit with TDO5-C-labeled target cells (Ramos), which can be attributed to the catalytic effect of TDO5-C on the target cell membranes. As a control, the SOG triggered by TDO5-C-labeled CCRF-CEM cells was studied. Little SOSG fluorescence enhancement was observed compared to the circuit only ( $A_1$ ,  $A_2$  and  $R_{12}$ ), indicating that the selective binding of TDO5-C can induce selective SOG.

### **Selective Cytotoxicity Effect to Cancer Cells**

Cell destruction by PDT was studied by irradiation with white light. As indicated by the SOSG studies,  $^1O_2$  is produced after the binding of TDO5-C with Ramos followed by the hybridization reaction of  $A_1$  and  $A_2$  to form  $A_{12}$  and the displacement of  $R_1$  from  $R_{12}$  by  $A_{12}$ . Therefore, the phototoxicity of the cell-surface circuit to cancer cells was studied by MTS [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium] assay. After 3 hr of light irradiation, the target cells (Ramos) and the control cells (CCRF-CEM) were cultured for 36 hr before evaluating cell viability with MTS reagent. Figure 2-8 shows the MTS data expressed as the mean viability (standard deviation). The statistical differences were assessed by Student's *t*-test. When  $A_1$ ,  $A_2$  and  $R_{12}$  were combined with nonlabeled cells and irradiated with white light, very little damage was observed to either target or control cells (cell viability of around 85%). These data are consistent with the previous SOG resulting from the weak leakage hybridizations of  $A_1$  and  $A_2$ . However, when 100 nM free C was incubated

with both cell lines, no statistical difference was evident ( $P > 0.30$ ), and both cell types showed low cell viabilities of about 50%. Herein, because the free C sequence did not selectively recognize target cells, it catalyzed the amplification reaction without selectivity and caused relatively equal cell death for both target cells and control cells. On the other hand, when each of the cell lines was first incubated with TDO5-C, rinsed, and then incubated with  $A_1$ ,  $A_2$  and  $R_{12}$ , high phototoxicity was observed for the target Ramos cells (45%), compared to 80% for the control cells (CCEF-CEM) with  $P < 0.02$ , indicating that TDO5-C can catalyze the hairpin circuit on specific cell membranes. In addition, we also found that the statistical toxicity difference between ( $A_1+A_2 + C+ R_{12}$ ) group and ( $A_1+A_2 + TDO5-C+ R_{12}$ ) group is evident ( $P < 0.001$ ) for CEM cells, but not for Ramos cells ( $P > 0.3$ ) as the consequence of catalytic selectivity of TDO5-C. Finally, to compare the cytotoxicity of our method with the 1:1 displacement method, we incubated the preannealed  $A_{12}$  (1  $\mu$ M) and  $R_{12}$  with cells and found a much higher cell viability of 79% for both cell lines. Under these conditions, the 1:1 displacement method did not show any selectivity to the two cell lines based on the lack of recognition element ( $P > 0.30$ ). These comparisons demonstrate the selective and amplified therapeutic effect of our method.

## Conclusions

In conclusion, our results demonstrate the feasibility of assembling a DNA circuit on cell membranes to achieve amplified and targeted photodynamic therapy. The DNA circuit, composed of four functional modules ( $A_1$ ,  $A_2$ ,  $R_{12}$  and TDO5-C) totally made of DNA, can greatly amplify the singlet oxygen generation and selectively kill cancer cells. In particular, the DNA hairpin amplification circuit can be catalyzed by specifically designed nucleic acid sequences. Many nucleic acids, including messenger RNA,

microRNAs and small interfering RNAs, are important biomarkers for various diseases.<sup>86-88</sup> If sequences for these biomolecules are available, the amplification hairpin DNA circuit can be designed to perform other biological and biomedical functions inside targeted disease cells with effective delivery methods. Second, increasing numbers of aptamers have been developed to target the membranes of a variety of cancer cell lines, thus establishing the universality of this DNA hairpin circuit for targeted and amplified therapy. Finally, as an application of DNA circuit to biological cells, the prototype DNA circuit demonstrated here has the potential to enhance DNA technology with new insights and will broaden the utility of DNA circuits for applications in biology, biotechnology, and biomedicine.

Table 2-1. Sequence of oligonucleotides used in this chapter. LNA bases are indicated by bold and underscores.

	Sequences (5'-3')
C (c*b*a*)	CGACATCTAACCTAGCTCACTGAC
A <sub>1</sub> (abcd*c*b*e*)	GTCAGTGAGCTAGGTTAGATGTCGCCATGTGTAGACGAC ATCTAACCTAGCACTTGTTCATAGAGCAC
A <sub>2</sub> (cdc*b*d*)	AGATGTCGTCTACACATGGCGACATCTAACCTAGCCCAT GTGTAGA
R <sub>1</sub> (eb)	Ce6 (FAM) GTGCTCTATGACAAGTGCTAGGTT
R <sub>1</sub> -LNA (eb)	Ce6 (FAM) <u>GTGCTCTAT</u> GACAAGTGCTAGGTT
R <sub>2</sub> (b)	ACTTGTCATAGAGCAC BHQ2 (DABCYL)
TDO5	AACACCGTGGAGGATAGTTCGGTGGCTGTTCAGGGTCTC CTCCCGGTG
TDO5-C	CGACATCTAACCTAGCTCACTGACTTTTTTTTTTTTTTTTTT AACACCGTGGAGGATAGTTCGGTGGCTGTTCAGGGTCTC CTCCCG GTG

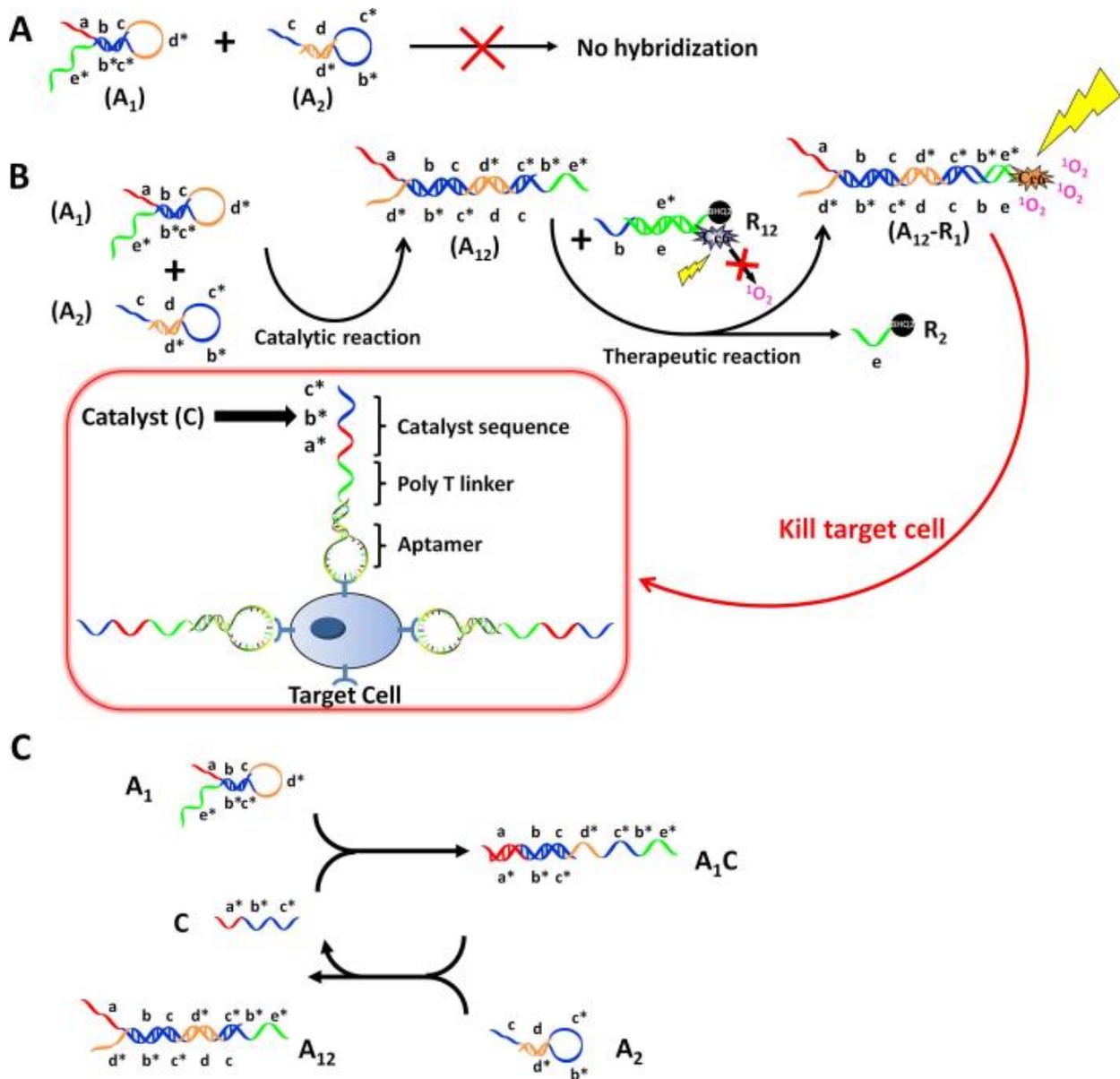


Figure 2-1. Illustration of DNA aptamer circuit on cell membrane. A). Scheme of the circuit without catalyst. B). Scheme of the circuit on cell membrane. The circuit involves two individual steps. In the catalytic step, target cell labeled with Apt -C catalyzes DNA hairpins  $A_1$  and  $A_2$  to form duplex  $A_{12}$ . In the therapeutic step,  $A_{12}$  can open duplex  $R_{12}$  and displace quencher-labeled single strand  $R_2$  to form  $A_{12}-R_1$ . Subsequently, Ce6-labeled  $R_1$  generates singlet oxygen ( $^1O_2$ ) to kill cancer cells by irradiation at 404 nm. C). Scheme of detailed reaction of DNA hairpins  $A_1$  and  $A_2$  catalyzed by C sequence. Different domains are labeled with different colors. All x domains are complementary to  $x^*$ .

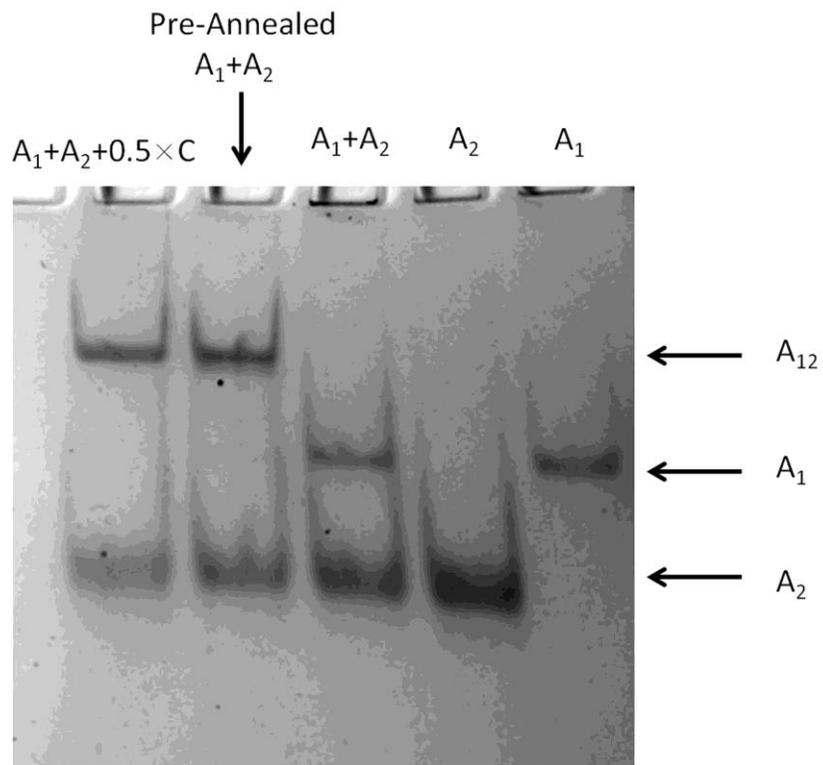


Figure 2-2. Image of the PAGE gel proving the catalytic effect of C sequence.

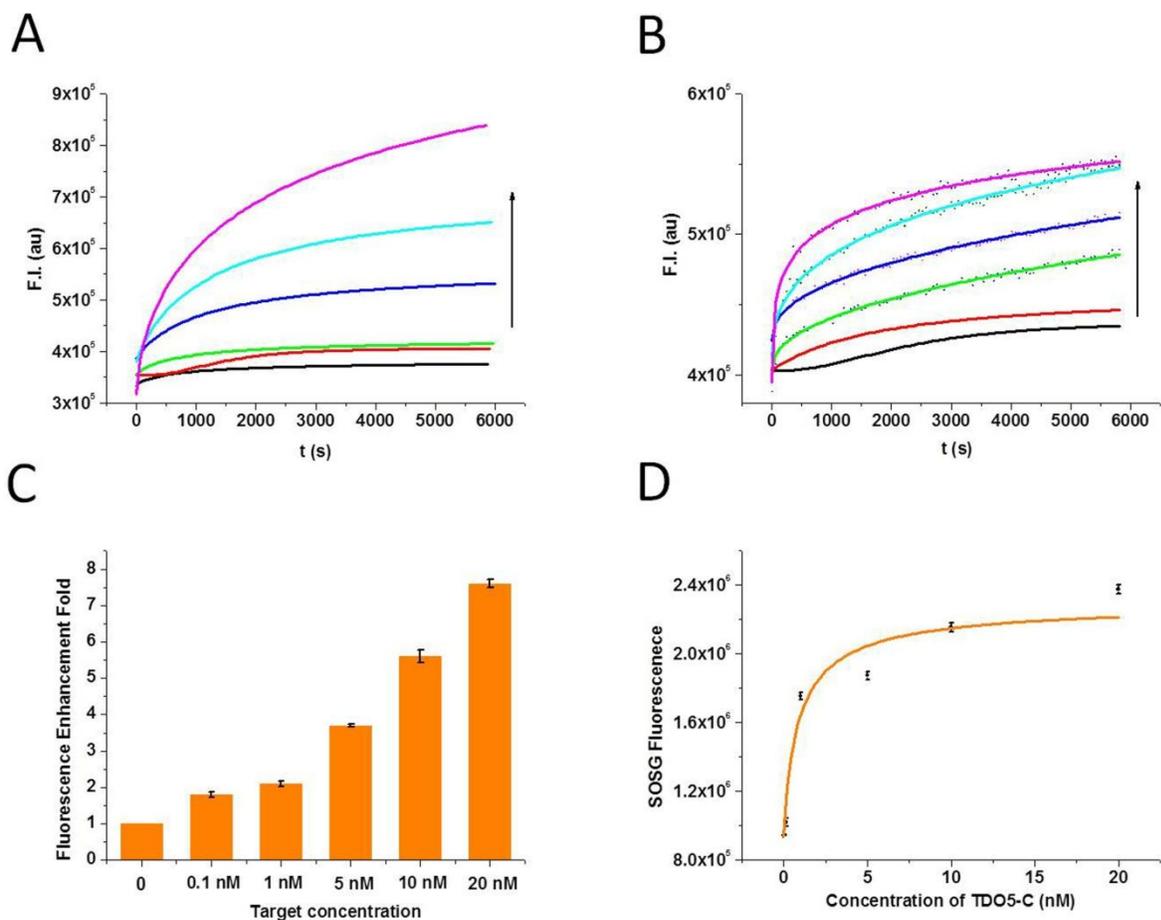


Figure 2-3. Signal amplification effect of the aptamer circuit in buffer. A). Kinetics of DNA circuit containing  $A_1$ ,  $A_2$  and  $R_{12}$  with different concentrations of TDO5-C (fluorescence intensities corresponding to  $F_{1:n}$ ) monitored by FAM fluorescence. The colored lines represent 0, 0.1 nM, 1 nM, 5 nM, 10 nM and 20 nM of TDO5-C, respectively. B). Kinetics of dsDNA  $R_{12}$  with different concentrations of  $A_{12}$  (fluorescence intensities corresponding to  $F_{1:1}$ ) monitored by FAM fluorescence. The colored lines represent 0, 0.1 nM, 1 nM, 5 nM, 10 nM and 20 nM of  $A_{12}$ , respectively. C). Comparison of the fluorescence enhancement fold of the catalytic DNA circuit and 1:1 displacement. The calculation is based on the equation  $\text{Fold} = (F_{1:n} - B) / (F_{1:1} - B)$ , where fold is the fluorescence enhancement ratio of 1:n method to 1:1 method, B is the background fluorescence. The fluorescence intensities at 6000 s are used to plot against different target concentrations. Each bar presents the mean and standard deviation derived from three independent experiments. D). The SOSG signal plotted as the function of TDO5-C concentration. The SOSG was triggered by irradiation at 404 nm, the maximum absorption of Ce6, for 10 minutes. The SOSG fluorescence was obtained with excitation at 494 nm and emission from 500 nm to 600 nm. Each bar presents the mean and standard deviation derived from three independent experiments.

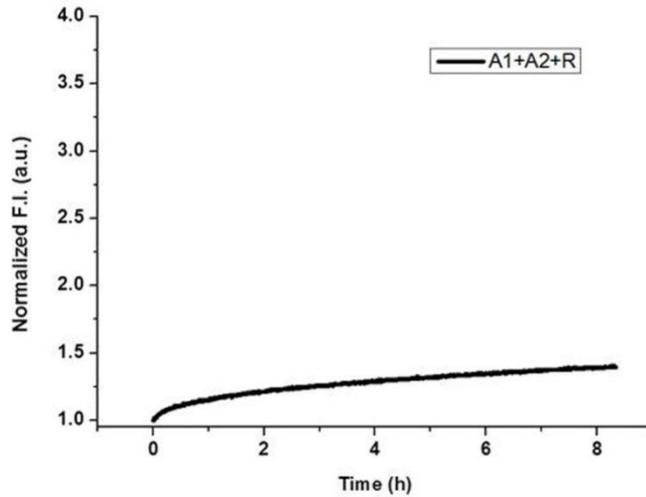


Figure 2-4. Fluorescence kinetics describing the leakage reaction of  $A_1$ ,  $A_2$  and  $R_{12}$  in the Fluo buffer. The data were normalized to the initial fluorescence intensity of the circuit.

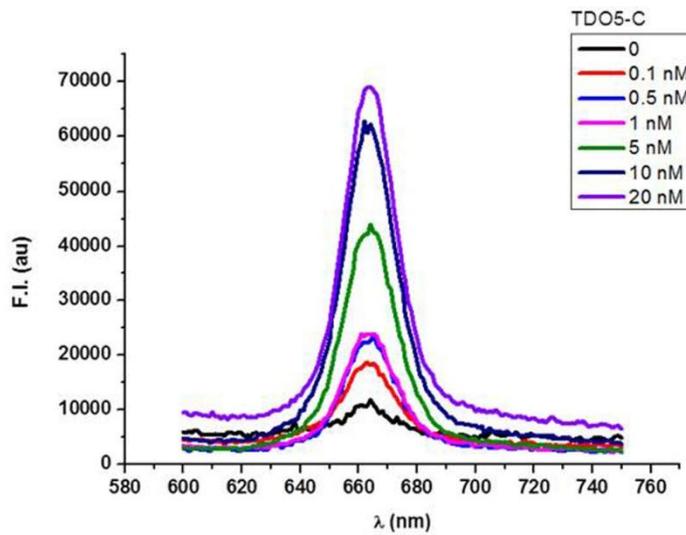


Figure 2-5. Fluorescence spectra of mixtures containing 100 nM  $A_1$ , 100 nM  $A_2$  and 150 nM Ce6-modified  $R_{12}$  with different concentrations of TDO5-C in buffer.

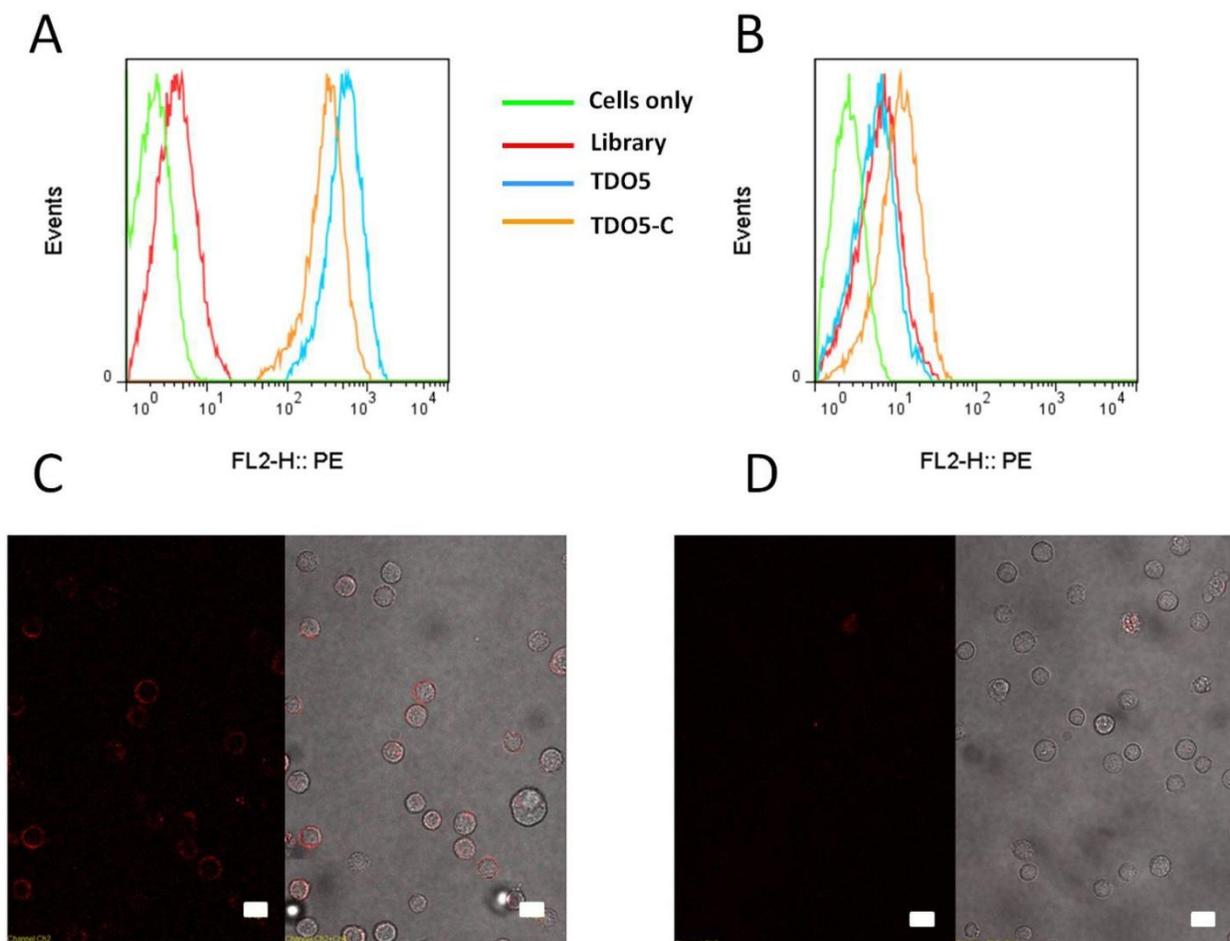


Figure 2-6. Demonstration of the selective recognition ability of the aptamer circuit. A) and B), Flow cytometry results of FAM-labeled TDO5-C binding with Ramos and CCRF-CEM. Aptamer TDO5 was used as positive control to show the maximum binding affinity. Cells: 200 k/sample; aptamer concentration: 250 nM. C) and D), Confocal imaging of Ramos cells and CCRF-CEM cells incubated with 250 nM TMR-modified TDO5-C at 4°C. Fluorescence image (left). Overlap of optical image and fluorescence image (right). The scale bar is 10  $\mu\text{m}$ .

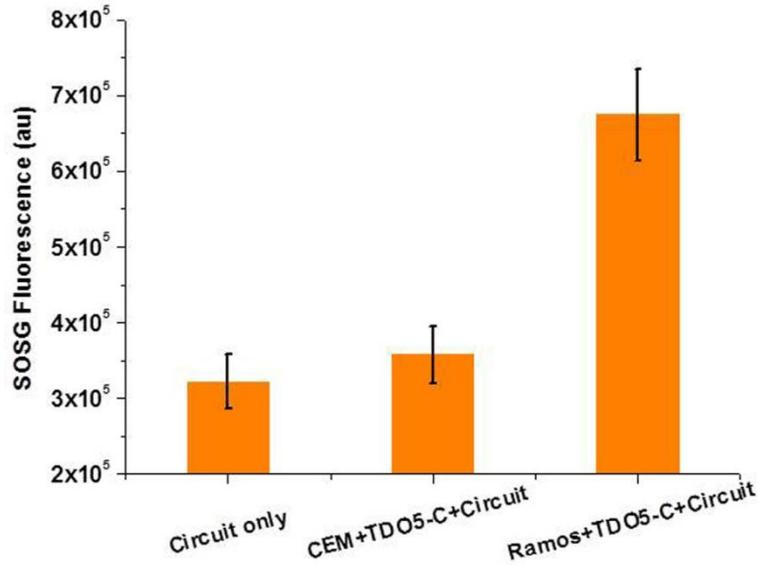


Figure 2-7. SOSG fluorescence of DNA circuit ( $A_1$ ,  $A_2$  and  $R_{12}$ ) incubated with buffer, TDO5-C-labeled CCEF-CEM cells (control) and TDO5-C-labeled Ramos cells (target). Each bar presents the mean and standard deviation derived from three independent experiments. (Cells: 200k/sample;  $\lambda_{ex}$ =494 nm and  $\lambda_{em}$ =532 nm).

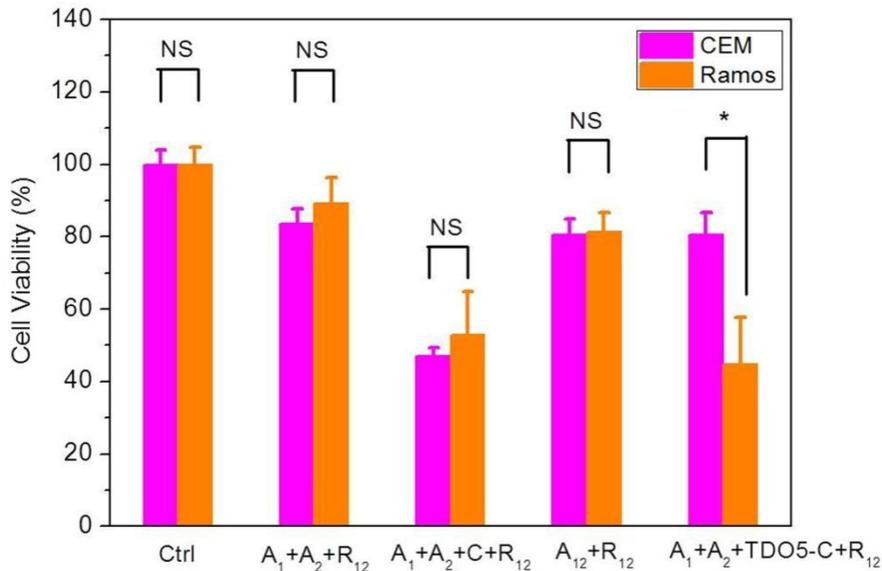


Figure 2-8. Cell viability result using MTS assay. The *in vitro* cytotoxicity was measured after 36 hr of incubation in cell medium with 3 hr of white light irradiation. Cells: 500k/sample. Each bar presents the mean and standard deviation derived from three independent experiments. P values were calculated by Student's *t*-test: ns, nonsignificance:  $P > 0.30$  and \* for  $P < 0.02$ ,  $n = 3$ .

### CHAPTER 3

## MOLECULAR ENGINEERING OF PHOTORESPONSIVE THREE-DIMENSIONAL DNA NANOSTRUCTURES FOR INTELLIGENT DRUG DELIVERY SIGNIFICANCE AND BACKGROUND

DNA is a useful construction material for various nanostructures by virtue of the remarkable hybridization specificity between complementary strands. Research on DNA nanostructures has not only improved the understanding of fundamental problems in genetics, but has also generated interest in exploring practical applications.<sup>89</sup> By using smart design, different static DNA nanostructures<sup>90,91</sup> have already been prepared.

Recently, reconfigurable DNA nanostructures capable of three-dimensional movement have attracted increasing interest because of their potential applications in intelligent drug delivery<sup>92</sup> and smart molecular sensing.<sup>93</sup> Active control of three-dimensional movement for DNA structures is usually achieved by input of specific molecular signals, such as DNA strands,<sup>47, 94</sup> enzymes<sup>95</sup> or protons<sup>96</sup>, to trigger a change in the shape or size of the structure. However, a major shortcoming arises from those designs because the accumulated DNA output waste can quickly deteriorate device performance and ultimately bring it to a halt.<sup>97</sup> To overcome this problem, a recycled drive power input is highly desired. Compared with other input signals, photons have significant advantages, such as clean, permanently high efficiency and no waste accumulation. In addition, by using light, DNA nanostructures can be remotely controlled, opening novel avenues in nanomedicine. Consequently, a photon-regulated, shape-changing DNA nanostructure would greatly contribute to applications in many fields of nanoscience.

Azobenzene has proven to be an effective photo-sensitive component because of its reversible stereo isomerization from the *trans* to *cis* forms at 300-380 nm and from

*cis* to *trans* at wavelengths >400 nm.<sup>98</sup> By incorporating azobenzene moieties into DNA double-stranded structures, hybridization can be controlled by the interconversion of azobenzene between the planar *trans* and nonplanar *cis* forms, allowing the formation of photocontrollable DNA structures.<sup>99, 100</sup> Here, we construct a photocontrollable, reconfigurable three-dimensional DNA tetrahedral cage using azobenzene incorporated- DNA.<sup>101-103</sup> By controlling the size of this cage, the release of encapsulated cargos (such as proteins or other macromolecules) can be triggered by light to produce a smart drug delivery system with precise temporal and spatial resolutions. In addition, the strategy of using azobenzene is universal and can be extended to any type of DNA nanostructure, providing new routes for manipulation of nanoscale shapes using photons.

A previously reported method was used to construct robust DNA tetrahedral structures by assembly of appropriately designed oligonucleotide sequences.<sup>6,37</sup> Inspired by the contraction and extension of a string, a hairpin structure was incorporated into the DNA tetrahedron. The hairpin can be opened and closed via hybridization and dehybridization, as shown in Figure 3-1. DNA sequences S1, S2, S3 and S4 form the contracted DNA tetrahedron with the hairpin in the closed state. Strands S5 with incorporated azobenzene moieties can hybridize with the hairpin portion, allowing the control of open-closed cycles of the hairpin structure by using UV and visible light. To achieve optimal photon-control, four S5 sequences were designed, as shown in Figure 3-1. The smallest number of azo moieties is 7 (S5-7Azo), with azobenzene inserted every three bases. Sequences with azobenzenes every two bases (S5-9Azo) and every one base (S5-17Azo) were also synthesized. Considering that the

distance of the azobenzene moieties from the sequence ends may influence hybridization efficiency, S5-10Azo was designed with a different azo-moiety distance from the sequence ends compared to S5-9Azo. With UV irradiation, the azobenzene molecules change to the nonplanar *cis* form separating strands S1 and S5, thus forcing the hairpin to form and finally converting the entire structure to the contracted state. On the other hand, when visible irradiation (>450nm) is applied, the azobenzene molecules convert back to the planar *trans* form and rehybridize with strand S1 to open the hairpin and extend the DNA tetrahedron. In this way, the shape of the entire DNA three-dimensional structure can be precisely controlled by photon irradiation.

## **Experimental Materials and Methods**

### **Chemicals and Regents**

The chemicals for synthesis of the phosphoramidite monomer and reagents for DNA modification were purchased from ChemGene (MA). The CPG columns for DNA synthesis were purchased from Glen Research (VA). The reagents for BIDBE, azobenzene phosphoramidite synthesis and gold nanoparticle synthesis were purchased from Sigma-Aldrich (MO).

### **Synthesis of DNA Sequences**

The DNA sequences were synthesized on the ABI 3400 DNA synthesizer. The synthesis protocol was set up according to the requirements specified by the reagents' manufacturers. Following on machine synthesis, the DNA products were deprotected and cleaved from CPG by incubating with 2.5 mL AMA (ammonium hydroxide/ Methylamine 50:50) for 17 hours at 40°C in water bath. The cleaved DNA product was transferred into a 15 mL centrifuge tube and mixed with 250 µL 3.0 M NaCl and 5.0 mL ethanol, after which the sample was placed into a freezer at -20 °C for ethanol

precipitation. Afterwards, the DNA product was spun at 4000 rpm under 3°C for 20 minutes. The supernatant was removed, and the precipitated DNA product was dissolved in 500 µL 0.2 M triethylamine acetate (TEAA Glen Research Corp.) for HPLC purification. The HPLC purification was performed with a cleaned Alltech C18 column on a Varian Prostar HPLC machine. The collected DNA product was dried and processed detritylation by dissolved and incubated in 200 µL 80% acetic acid for 20 minutes. The detritylation DNA product was mixed with 400 µL ethanol and dried by a vacuum dryer. The DNA products were quantified and stored in DNA water for subsequent experiments. The detailed sequences information are described in

### **Synthesis of BF-PS-DNA**

For the synthesis of BF-PS-DNA, the method reported by Lee et al.<sup>104</sup> was used. First, phosphorothioate-modified DNA (PS-DNA) was synthesized at a specific position on DNA strand 1 using sulfuration reagent (Glen Research). The ligand BIDBE was then synthesized by using the protocol reported by Luduena et al.<sup>105</sup> DNA, BIDBE solution, and Tris-HCl buffer were incubated at 50°C for 5-6h to form BIDBE-DNA. The best ratio for BIDBE and DNA phosphorothioate sites was around 200:1. Then, to reduce the disulfide bond, 100uL 10uM BIDBE-DNA solution, 1uL 50mM acetate buffer and 10uL 1mM TECP were mixed and incubated at room temperature for 2 hours. Gel-filtration was used to remove the impurities.

### **Assembly of DNA Tetrahedral Nanostructures**

Assembly of contracted DNA tetrahedra: 10uL 10uM DNA strand S1, 10uL 10uM DNA strand S2, 10uL 10uM DNA strand S3, 10uL 10uM DNA strand S4, 10uL 150mM Tris-HCl (pH=7.5), 10uL 150mM MgCl<sub>2</sub> and 40uL DNA water were mixed and annealed at 95°C for 5min and cooled to room temperature in approximately 5 minutes.

Assembly of extended DNA tetrahedra (no azobenzene incorporated): 10uL 10uM DNA strand S1, 10uL 10uM DNA strand S2, 10uL 10uM DNA strand S3, 10uL 10uM DNA strand S4, 10uL S5-cDNA (no azobenzenes), 10uL 150mM Tris-HCl (pH=7.5), 10uL 150mM MgCl<sub>2</sub> and 40uL DNA water were mixed and annealed at 95°C for 5 min and cooled to room temperature in approximately 5 minutes.

Assembly of azobenzene-incorporated DNA tetrahedra: 10uL 10uM DNA strand S1, 10uL 10uM DNA strand S2, 10uL 10uM DNA strand S3, 10uL 10uM DNA strand S4, 10uL DNA strand S5, 10uL 150mM Tris-HCl, 10uL 150mM MgCl<sub>2</sub> and 30uL DNA water were mixed and annealed at 95°C for 5min and cooled to room temperature in approximately 5 minutes. Tetrahedra were purified using PAGE gel.

### **Native Polyacrylamide Gel Electrophoresis (PAGE) for Structural Characterization**

Polyacrylamide solution (40%, 750uL) was diluted to 6% by adding 4.25mL 10mM TAE buffer containing 15mM MgCl<sub>2</sub>. Fifty uL APS and 5uL TEMED were added to the polyacrylamide solution to polymerize it. After loading 10uL 1uM DNA sample solutions, the gels were run on an electrophoresis unit (Biorad) at 4 °C using a constant voltage of 80 V for 90 minutes. After electrophoresis, the gels were stained with Stains-All for 30 min and imaged using a digital camera.

### **Phosphorothioate DNA Tetrahedron Assembled with Gold Nanoparticles**

It has been reported that AuNPs are more stable when capped with phosphine reagent.<sup>106</sup> Therefore, in this step, 10mL 60nM gold nanoparticle solution was mixed with 2mg bis(para-sulfonatophenyl) phenylphosphine dihydrate dipotassium salt (Strem Chemicals, Newburyport, MA) and shaken at 25 °C overnight. After precipitation and centrifugation, the supernatant was removed, and the AuNP precipitate was redispersed into water. After stabilization, AuNPs were assembled with the DNA tetrahedron to

visualize the shape change of structures. Ten  $\mu\text{L}$  100nM DNA tetrahedron solution was mixed with 50 $\mu\text{L}$  60nM AuNPs and incubated overnight at room temperature. The best TEM images were obtained from the mixture solution containing the tetrahedra and AuNPs at a ratio of 1:3.

### **FRET Measurement of Structure Change in Response to UV and Visible Irradiation**

After assembly, DNA tetrahedral structures were excited at  $\lambda_{\text{exc}}=488\text{nm}$  at constant temperature. After that, the light source in the fluorometer was used to irradiate the solution at 350nm for 3 min, and the fluorescence spectrum was obtained again immediately after the UV irradiation. Next, visible light (450nm) was applied to the structures for another 3 min, and fluorescence spectrum of the solution was measured. Additional cycles with alternate UV and Visible irradiation were performed, and the fluorescence spectra were recorded.

### **AFM and TEM Measurements**

AFM experiments were carried out on a Nanoscope IIIa (Veeco, Santa Barbara, CA) using tapping mode in ambient air. The radius of curvature of silicone tip was about 10 nm. All topographic images were obtained with 512 x 512 pixels<sup>2</sup> at a scan rate of 1.5 Hz. After annealing strands S1, S2, S3, S4 and S5-10Azo under visible light, the structures were scanned by using AFM. The tetrahedra were expected to have one 10.5nm edge and five 7nm edges for the extended state, and based on the AFM image, the sizes and heights were consistent with the calculated values.

Transmission electron microscopy (TEM) images were obtained on a Hitachi H-7000 NAR transmission electron microscope. For the extended structures, samples were dried on copper film under visible light. Then a parallel sample was irradiated at

350nm until it was completely dry on the copper film. Afterwards, the samples were imaged by using TEM at a working voltage of 100 kV.

## **Results and Discussion**

### **Construction and Optimization of DNA Tetrahedral Nanostructure**

The sequence design for assembly of tetrahedral structure will mainly follow the method of Turberfield et al.<sup>107</sup> A stable and pure tetrahedral structure must be confirmed before further experimentation by consideration of two criteria. First, the hairpin that will be incorporated into the tetrahedron should be able to be opened and closed via DNA hybridization and dehybridization. Normally, a 5-base stem and 12-14 base loop is the optimum design for rapidly responsive molecular beacons. Therefore, we will incorporate a hairpin with a 5-base stem and a 12-base loop into the large structure. Meanwhile, the numbers of azobenzene moieties will be optimized to achieve superior photon-controllability, as illustrated in Figure 3-1. The smallest number of azo moieties will be 7 (S5-7Azo), with azobenzene inserted every three bases. Sequences with azobenzenes every two bases (S5-9Azo) and every one base (S5-17Azo) will also be synthesized. Considering that the distance of the azobenzene moieties from the sequence ends may influence hybridization efficiency, S5-10Azo will be designed with a different azo-moiety distance from the sequence ends compared to S5-9Azo.

### **Structural Confirmation of DNA Tetrahedral Nanostructure**

Native polyacrylamide gel electrophoresis (PAGE) and atomic force microscopy (AFM) were used to confirm the formation of azobenzene-incorporated DNA tetrahedra. The gel result in Figure 3-2A shows that the azobenzene-incorporated tetrahedral structures (from lanes 3 to 6) can form stable extended structures under visible irradiation compared with T1 (Lane 1) and T2 (Lane 2). T1 and T2 are contracted and

extended tetrahedral structures without any azobenzene moieties. However, some contracted structures also appeared along with the extended forms, and they became more and more obvious as the number of azobenzene moieties increased, indicating that hybridization efficiency is affected by the azobenzene moieties. The four azo-incorporated structures have slightly lower mobilities than T2 because the azo-moieties increase the molecular weight of the structures compared with T2. Thus, in the presence of azobenzene-incorporated strands, the DNA tetrahedral structures can still be successfully constructed.

AFM images were taken on a Nanoscope IIIa atomic force microscopy using the tapping mode in ambient air. After annealing strands S1, S2, S3, S4 and S5-10Azo under visible light, the structures were scanned by AFM (Figure 3-2B). The tetrahedra were expected to have one 10.5nm edge and five 7nm edges for the extended state, and based on the AFM images, the sizes and heights were consistent with the calculated values.

### **Characterization of DNA tetrahedral structure with photocontrollability**

Fluorescence measurements were used to demonstrate the photo-control of the Azo-DNA tetrahedral structures. Fluorophores (FAM) and quenchers (Dabcyl) were incorporated on the two ends of the hairpins in S1. When the tetrahedron is in the contracted state, the fluorescence is quenched as a result of the close proximity of FAM and Dabcyl. When the hairpin is opened to form the extended structure, the fluorescence intensity should increase by the separation of fluorophore and quencher. Fluorescence signals of structures T1 and T2 were used to obtain the background and maximum signals, respectively. In Figure 3-3A, when UV irradiation was applied to T-10Azo, the fluorescence intensity decreased because the dehybridization of azo-

incorporated S5 strands promoted the formation of hairpin structures. Different azobenzene-incorporated structures were tested at 48°C, and their fluorescence signals were normalized based on the maximum signal of T2 (Figure 3-3B). The results showed that tetrahedral structures containing a low percentage of azobenzenes (T-Azo7) did not respond differentially to either UV or visible light, indicating that S5 did not dissociate effectively from the structures. On the other hand, structures possessing the maximum number of azobenzenes (e.g., T-Azo17) showed poor hybridization between S5 and S1, even under visible irradiation. Based on these findings, we concluded that the structures containing 1 azobenzene between every two nucleotides in S5 (i.e., T-Azo9 and T-Azo10) give the best photon response efficiency and optimum structural stability.

Figure 3-3C shows the fluorescence changes as the T-10Azo structures extended and contracted with successive irradiation by UV (3 minutes) and visible light (3 minutes). The efficiency did not decrease, even after 10 open/closed iterations with no addition of extra oligonucleotides and no generation of waste strands. Therefore, this photon-regulated DNA structure is stable and robust. Additionally, all the cycles were performed at 48°C in order to have rapid dynamic response to the wavelength changes. Here, the melting temperatures of S5-10Azo and S1 in the *cis* and *trans* configurations are 30.1°C and 55.7°C, respectively, and the photon-fueled reconfigurable structures work most efficiently at around 45-55°C. The relatively high experimental temperature is the reason for the relatively high background signal. Overall, these results demonstrated successful, long-lasting control of structural alternation with photon energy input.

To visualize the structural changes when applying different wavelengths of light, gold nanoparticles were assembled with tetrahedral structures. Taking advantage of Au-

NPs' uniform size, clear visualization by TEM, and site-specific control on DNA strands by using phosphorothioate DNA and bifunctional linkers<sup>104</sup>, gold nanoparticles (diameter of 3.5nm) were attached on the three vertices of the variational triangular face of the tetrahedral structures (Figure 3-4). When the shapes of the tetrahedron change, the relative positions of particles on the vertices of the triangle change accordingly. This method allowed clear visualization of the movement of DNA structures triggered by light. In Figure 3-5A, the sizes of extended structures matched well with the calculated values. The two isosceles edges are around 7nm, and the bottom edge is 11nm long. After UV irradiation, some tetrahedra converted to contracted structures, causing the bottom edges of triangles to shrink to 4nm, as shown in Figure 3-5B.

To remove the interference of random alignment of AuNPs, control experiments were also performed. For the DNA tetrahedral structures without any phosphorothioate modification, no regular triangular structures were found on the TEM image (Figure 3-5C). These images were consistent with the fluorescence data and demonstrated the successful working mechanism of our photon-regulated reconfigurable DNA nanostructures.

## **Conclusions**

In summary, we have successfully used azobenzenes to construct a tetrahedral DNA nanostructure controlled by photons. PAGE and AFM were utilized to confirm the existence of the tetrahedral DNA nanostructures. Fluorescence intensity changes, as well as AuNP-assisted TEM measurements, demonstrated the photon-controllability. These results indicate that incorporation of azobenzene moieties into DNA strands allows control of three-dimensional structure, in line with our previous report which demonstrated the control of a two-dimensional DNA hairpin structure.<sup>103</sup> We believe that

this photo-responsive nanocage will greatly facilitate the development of DNA structures as drug delivery platforms for intelligent therapy. For example, one potential application is the utilization of the 3-D structural changes to trigger the release of cargos (such as proteins or other macromolecules) encapsulated in the DNA nanostructure, as a smart drug delivery system with precise temporal and spatial resolution.

Table 3-1. DNA sequences used for tetrahedron assembly

	Sequence (5'-3')
S1	GGT GAT AAA ACG TGT AGC AAG CTG TAA TCG ACT CTA Dabcyl GGC GGA AGA ACC CAC AAC CGC C FAM CGC TCA CTA CTA TGG CG
S2	AGG CAG TTG AGA CGA ACA TTC CTA AGT CTG AAA TTT ATC ACC CGC CAT AGT AGA CGT ATC ACC
S3	CTT GCT ACA CGA TTC AGA CTT AGG AAT GTT CGA CAT GCG AGG GTC CAA TAC CGA CGA TTA CAG
S4	ATCTAACTGCTGCGCCGCCGGGAAAATACTGTACGGTTAGATTTTT TTTTT
S5-cDNA	GCG GTT GTG GGT TCT TCC GC
S5-7azo	GXC GGX TTG XTG GXG TTX CTT XCC GXC
S5-9azo	GCX GGX TTX GTX GGX GTX TCX TTX CCX GC
S5-10azo	GXC GXG TXT GXT GXG GXT TXC TXT CXG GXC
S5-17azo	GCX GXG XTX TXG XTX GXG XGX TXT XCX TXT XCX CXG C

X represents azobenzene moiety

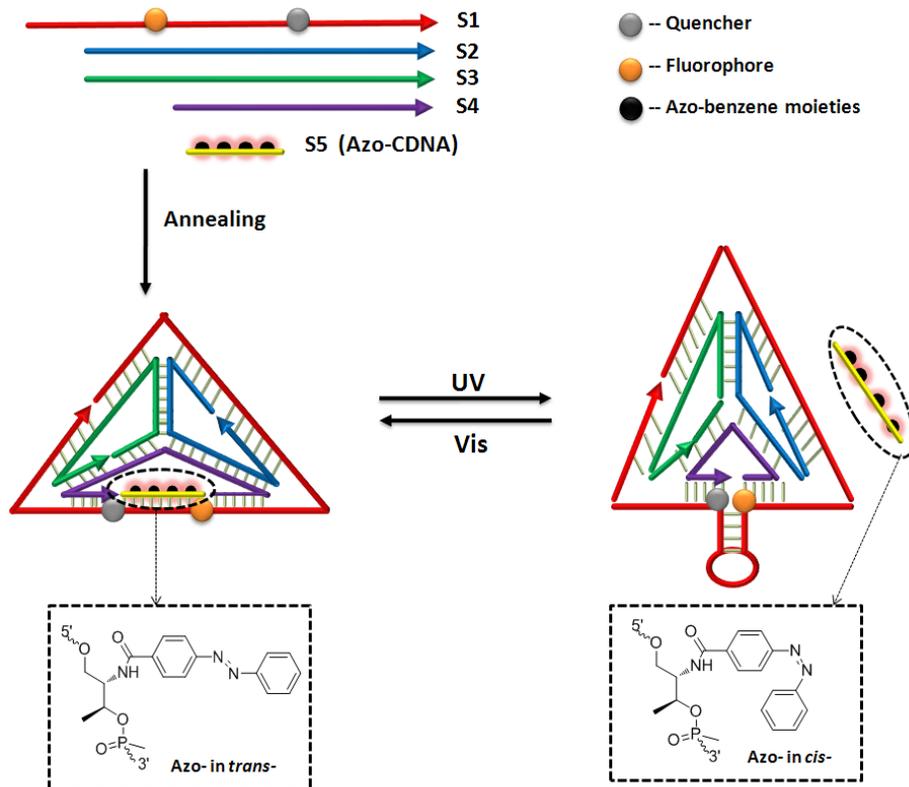


Figure 3-1. Design of photocontrollable DNA nanostructure using azobenzenes. S1, S2, S3, S4 and S5 were mixed in the ratio of 1:1:1:1:1 to form the azo-incorporated tetrahedral structures. Four S5 sequences were designed with different azobenzene moiety numbers. S5-7Azo, S5-9Azo, S5-10Azo and S5-17Azo contain 7, 9, 10 and 17 azobenzene moieties, respectively. A fluorophore and a quencher were incorporated at the ends of the hairpin to indicate structural movement.

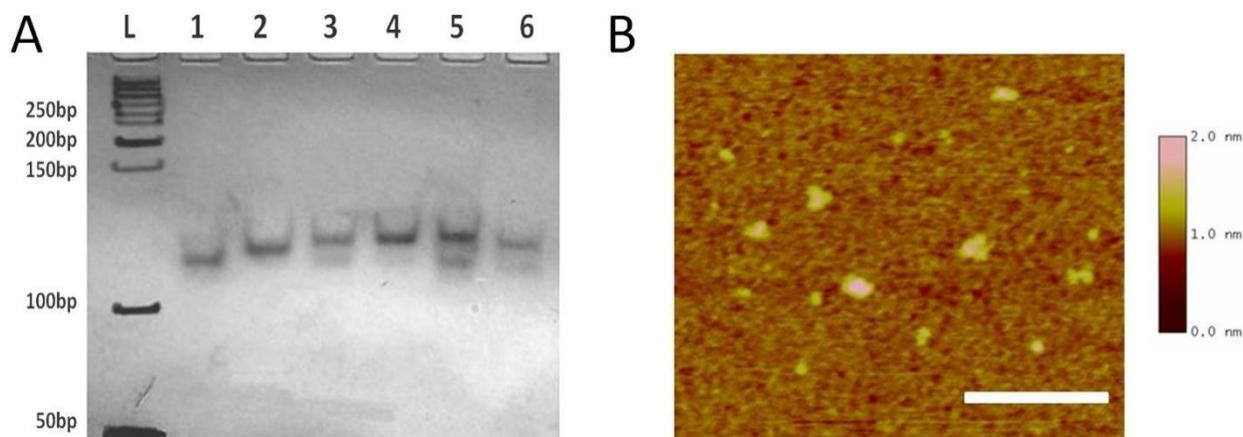


Figure 3-2. A) Native polyacrylamide gel electrophoresis (6%) analysis of different azo-incorporated tetrahedral structures at 4 °C. Lane 1 corresponds to the contracted tetrahedron without S5. Lane 2 is the extended tetrahedron with S5-cDNA (no azobenzenes). Lanes 3-6 are the bands of T-7Azo, T-9Azo, T-10Azo and T-17Azo, respectively. L: 50-bp ladder consisting of double strands of DNA with length increase in 50-bp steps. B) AFM image of T-10Azo structures recorded with 10nm tips. (Bar length: 100nm).

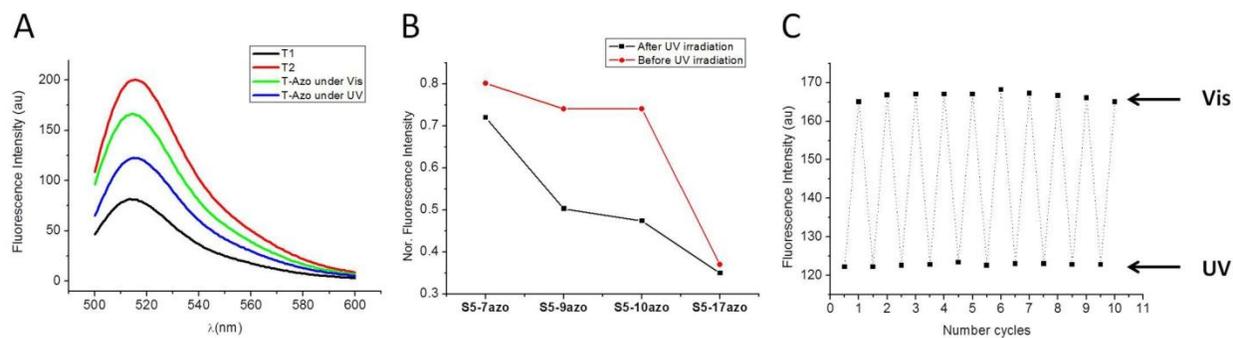


Figure 3-3. Fluorescent characterization of DNA tetrahedral structure with photocontrollability (A). Fluorescence measurement of T-10Azo labeled with FAM and Dabcyl in response to UV irradiation. (B).Fluorescence signal differences in response to UV irradiation for different azo-incorporated tetrahedral structures. (C). Cycling of the closed-open forms at 48°C by repeated visible and UV irradiations.

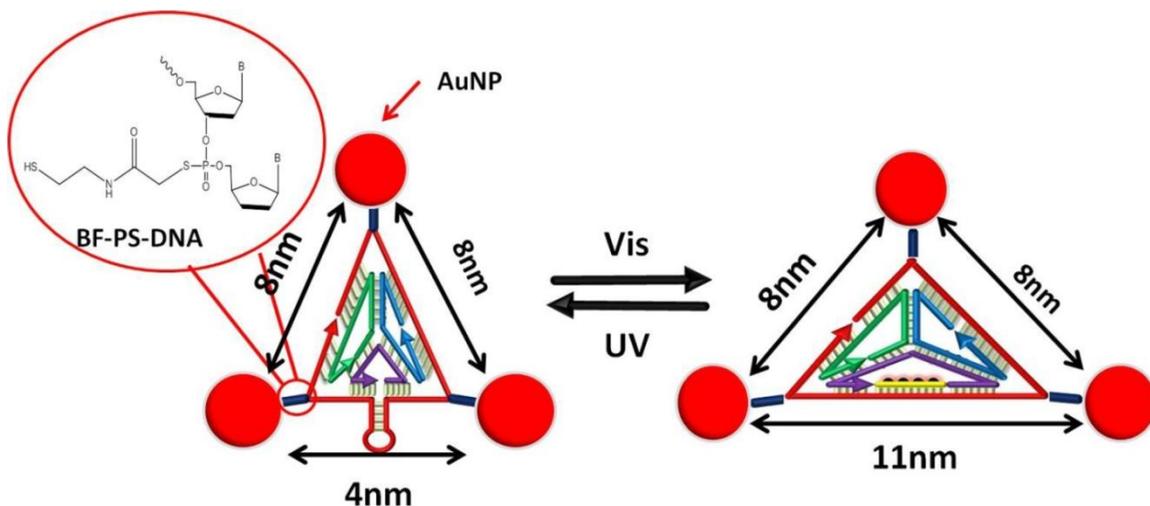


Figure 3-4. Illustration of gold nanoparticle (AuNP) assembly on three vertices of the variational triangular faces of tetrahedra by a phosphorothioate anchor and a short bifunctional fastener (BF). AuNPs can be attached to the BF-PS-DNA sites to observe different triangular shapes under UV and visible irradiation.

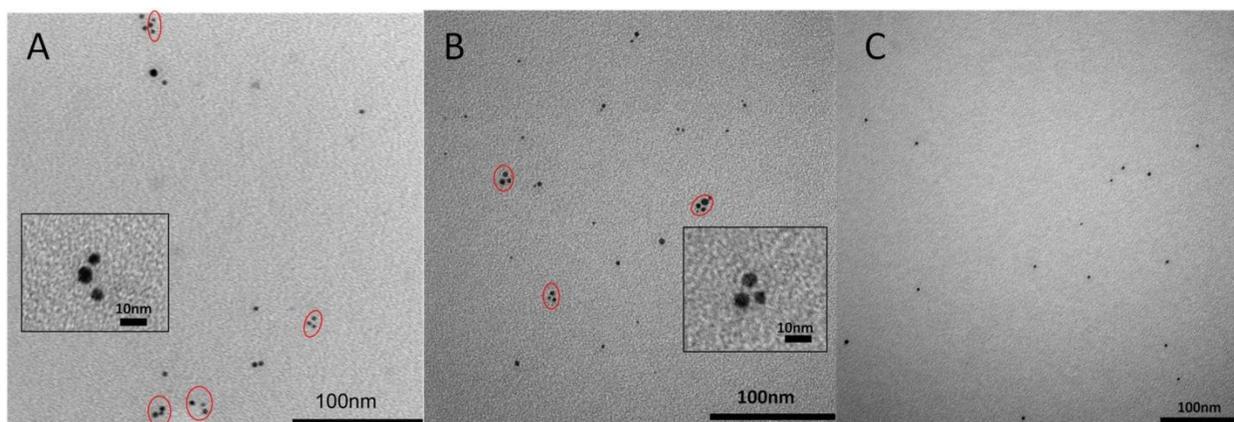


Figure 3-5. TEM images of AuNPs assembled on three vertices of the variational triangular face of tetrahedra. A) Before UV irradiation. B) After UV irradiation. C) Tetrahedra without any phosphorothioate modifications are shown as control.

## CHAPTER 4 MOLECULAR ENGINEERING OF A DNA LOGICAL CIRCUIT FOR PROGRAMMABLE AND AUTONOMOUS REGULATION OF PROTEIN ACTIVITY

### **Significance and Background**

Nucleic acids, as carriers of genetic information with well-regulated and predictable structures, are promising materials for the design and engineering of biochemical circuits. In particular, by emulating the digital logic found in typical circuit boards, recent reports demonstrated that DNA-based biocircuitry could perform logic gate operations, signal restoration, amplification, feedback, and cascading, all by distinct DNA strands.<sup>49, 51, 93, 108, 109</sup> Recently, their ability to interact with naturally occurring biomolecules, together with such unique properties as programmability, Boolean processing capability, nanometric size, and autonomous operation, has opened a novel and exciting direction in biological and biomedical applications.<sup>110-115</sup> Various circuits with single-purpose or generic computing capability have been demonstrated using nucleic acid (NA) base pairing interactions. Examples include RNA-based logic devices for processing cellular information<sup>116</sup> and quantitatively programming gene expression,<sup>117</sup> as well as a DNA-based biocomputer for logical analysis of multiple mRNA to trigger corresponding cellular response.<sup>118</sup> Such devices demonstrated the programmability and versatility of NA circuits for further development toward the goal of logically analyzing a complex biological environment and precisely regulating the actuation of cellular behaviors. While these NA-based circuits are of great scientific interest, they are primarily based on nucleic acid hybridization and strand displacement reactions between NA probes of different lengths. This has severely restricted circuit operation solely to genetic molecules.

Inspired by these advances, but distinct from their operation at the level of gene expression by DNA hybridization, we sought to explore the interactions between NA and other molecules such as proteins in cell-free logical circuit operation. There are two significant issues of interest: the first is to understand whether NA-protein interactions can be programmably enrolled in molecular circuits; and the second is to apply the capability of the logic circuit to directly perform precise and smart manipulation of the function of proteins, which are often more important than genetic molecules because proteins are at the centers of homeostatic systems and the key regulators of an organism's behavior.

A major challenge for direct protein manipulation by DNA logic circuitry is to find a key component that can bridge nucleic acids with proteins without influencing the programmability and versatility of logic circuits. A common regulation paradigm used in previous reports is genetic regulation of protein expression, which requires sophisticated and cell-wide coordination.<sup>119-121</sup> However, a special single-stranded oligonucleotide, commonly known as an aptamer, has the potential to interact with proteins specifically and is thus explored and developed in this report as a simple and effective molecular circuit for manipulation of protein functions and activities. Aptamers, obtained via an *in vitro* selection strategy called systematic evolution of ligands by exponential enrichment (SELEX), can extend the recognition capabilities of nucleic acids from Watson-Crick base-pairing to interactions with various targets, such as small molecules, proteins, and even viruses or cells, via the aptamer's unique secondary or tertiary structure.<sup>10, 76, 77</sup> In addition, the affinity and specificity of an aptamer can be tuned through the selection process or by post-selection sequence optimization in order

to meet the specific performance requirements of a given application. Finally, some aptamers are not only able to recognize the target proteins but can also regulate protein functions.<sup>122-124</sup> Such aptamers are regarded as potential drugs with protein regulation functions or drug carriers for many diseases and are already in the pipeline for clinical use, including PDGF and VEGF aptamers for controlling the age-related macular degeneration,<sup>125, 126</sup> demonstrating their reliability for biomedical applications. Since an aptamer is essentially a single-stranded oligonucleotide, it is convenient in logic circuit design, just as previous NA-based circuits. Aptamers can be directly used as building blocks to fabricate seamless logic-based aptamer circuits with enhanced capabilities and an extended scope of applications, from simple DNA base-pairing reactions to more complicated biomolecular reactions such as NA-protein or NA-small molecule interactions.

As a proof of concept, human  $\alpha$ -thrombin, which initiates blood coagulation processes by converting fibrinogen to fibrin, was chosen as our model target protein. Imbalances in thrombin levels can lead to a variety of functional disorders, even death. For example, thrombin excesses can result in life-threatening blood clots in key organs.<sup>127</sup> Consequently, the design of safe and effective anticoagulants would have a significant effect on clinical therapy. However, dosage control, where imprecision can lead to severe side effects, is a challenge for traditional anticoagulants. Therefore, inspired by the negative feedback loop concept of biological signaling pathways, in which a signal bias induces the expression of its own inhibitor, we designed a logic-based molecular circuit that can precisely sense the local enzymatic environment (i.e., the concentration of thrombin in the present example) and smartly control its

coagulation function by a concentration-triggered threshold control module, which has the potential to provide a more efficient and safer therapeutic strategy.

The operating principle of our molecular circuit for protein regulation is schematically illustrated in Figure 4-1A. In general, a programmable and autonomous circuit with threshold control was constructed of three DNA modules: an Input Convertor that converts the protein input to DNA input for downstream cascade reactions; a Threshold Controller that sets the threshold concentration for the system to maintain regular protein activity; and an Inhibitor Generator that inhibits excessively high protein activity once it surpasses the threshold. This circuit can intelligently sense the activity, i.e., the concentration, of protein, and initiate the inhibitory function through a threshold control loop when excessively high protein activity occurs. By setting the threshold value according to each practical situation, the circuit may be usable as a smart drug delivery system in the design of personalized medicine.

To demonstrate such intelligent regulatory function with thrombin as a model, two anti-thrombin aptamers are employed to build an aptamer circuit to smartly control coagulation: a 29mer (TA-29) that binds to the heparin exosite without inhibitory function and a 15mer (TA-15) that binds to the fibrinogen exosite with strong inhibitory function.<sup>128, 129</sup> In the detailed design (Fig. 4-1B), the circuit includes a series of aptamer and DNA displacement reactions, in which a single-stranded DNA (ssDNA) can be displaced from the initial duplex by an even stronger binder, either a protein molecule or a better matched DNA strand.<sup>51, 130, 131</sup> The circuit starts with the introduction of thrombin. In the Input Convertor, thrombin reacts with duplex Aptamer-Input (A-I), which contains TA-29 partially hybridized with a piece of ssDNA. This ssDNA, termed as DNA-

input, is released from A-I by competitive binding of thrombin to TA-29, converting the protein input to DNA input for the following cascade reactions.

The DNA-input then enters the Threshold Controller module and rapidly reacts with duplex Threshold (T) via an exposed toehold  $a^*b^*$  to generate inert ssDNA *Waste* ( $W_1$ ) without further reaction. Through this bypass route, thrombin only binds with TA-29 and thus can still perform its normal catalytic function in blood coagulation. However, after T is depleted, the excess DNA-input will continue to the Inhibitor Generator module, in which DNA-Input reacts with duplex Output (O) via exposed toehold  $b^*$ , thereby triggering the amplification reaction of O with Fuel (F), i.e., signals that help to catalytically produce the output (Figure 4-2). The released product, which is denoted as S (ctd') with effective toehold t, then cascades to duplex Generator (G), followed by the release of the Inhibitor TA-15 to inhibit thrombin coagulation.

The sequential order of these reactions is precisely controlled by the differences in thermodynamic stability, as well as reaction kinetics, between each component. In this system, stability is mainly determined by the length of exposed toehold. For instance, both T and O can react with DNA-input, but the reaction of DNA-input with T is more favorable than its reaction with O, based on the longer toehold of T ( $a^*b^*$ ) compared to O ( $b^*$ ). As a result, more thermodynamically stable duplex strands will be formed between the DNA-input and T ( $a^*b^*c$ ). The difference in thermodynamic stability is also reflected in the reaction kinetics. In this case, the 10-nt toehold in T provides a displacement reaction rate about 100 times higher than that of the 5-nt toehold in O.<sup>130</sup> As a result, Inhibitor Generator can only work when the Threshold Controller is completely consumed.

An additional concern is the need for sufficient concentration of O to produce an adequate response. To ensure this reaction, an entropy-driven amplification strategy is introduced via the Fuel strand in the Inhibitor Generator. DNA-input can transform free Fuel into output without being consumed, according to the reactions shown in the Figure 4-2.<sup>51</sup> Thus, a small amount of DNA-input can trigger the release of a large amount of inhibitor (TA-15). Furthermore, to avoid undesired leaking reactions, the lengths of certain strands were shortened (e.g., d' is shorter than d). Overall, by precisely and smartly programming the major duplex elements A-I, T, O and G in the three modules, the aptamer circuit can work through an autonomous threshold control loop to generate Inhibitor TA-15 and intelligently regulate the activity of thrombin according to the preset threshold concentration.

## **Experimental Materials and Methods**

### **DNA Synthesis**

The DNA sequences were synthesized on the ABI 3400 DNA synthesizer. The synthesis protocol was set up according to the requirements specified by the reagents' manufacturers. Following on machine synthesis, the DNA products were deprotected and cleaved from CPG by incubating with 2.5 mL AMA (ammonium hydroxide/Methylamine 50:50) for 17 hours at 40°C in water bath. The cleaved DNA product was transferred into a 15 mL centrifuge tube and mixed with 250 µL 3.0 M NaCl and 5.0 mL ethanol, after which the sample was placed into a freezer at -20 °C for ethanol precipitation. Afterwards, the DNA product was spun at 4000 rpm under 3°C for 20 minutes. The supernatant was removed, and the precipitated DNA product was dissolved in 500 µL 0.2 M triethylamine acetate (TEAA Glen Research Corp.) for HPLC purification. The HPLC purification was performed with a cleaned Alltech C18 column on

a Varian Prostar HPLC machine. The collected DNA product was dried and processed detritylation by dissolved and incubated in 200  $\mu$ L 80% acetic acid for 20 minutes. The detritylation DNA product was mixed with 400  $\mu$ L ethanol and dried by a vacuum dryer. The DNA products were quantified and stored in DNA water for subsequent experiments. The detailed sequences information are described in supporting information (Table 4-1 and Table 4-2).

### **DNA Purification**

Native PAGE was applied to purify the Aptamer-Input, Threshold, Output and Generator duplex strands to remove excess strands and avoid undesired system leakage. The ssDNA components of A-I, T, O and G were annealed at concentrations of around 50  $\mu$ M in 1 $\times$  TAE-Mg buffer (40 mM Tris-Acetate-EDTA, pH 8.0, 12.5 mM  $\text{Mg}(\text{Ac})_2$ ). Native PAGE gels (12%) in 1 $\times$  TAE-Mg buffer were run at 110 V for 90 minutes at 4 $^\circ$ C and stained with GelRed stain solution (Biotium, CA). Only the sharp bands were cut from the gels, chopped into small pieces, and soaked in 1 $\times$  TAE-Mg buffer for 24 hours. After soaking out most DNA molecules from the gel pieces, the solutions were extracted and concentrated with centrifugal filter devices (Millipore, MA). Finally, the DNA duplex sequences were quantified by UV spectrometry and kept in buffer for future use.

### **Quantitative Analysis of Released DNA-input Strands Triggered by Thrombin**

A 100  $\mu$ L sample of 100 nM A-I probes was placed in thrombin buffer containing 1 $\times$ TAE with 100 mM NaCl, 12.5 mM  $\text{MgCl}_2$ , 10 mM KCl, and 1 mM  $\text{CaCl}_2$ . Different amounts of thrombin were incubated with the probes, and the fluorescence was monitored. To quantify the released DNA-input, 1 mL of 500 nM FAM-TA-29 and 1 mL of

500 nM DABCYL-DNA-input were annealed in thrombin buffer to make the FRET A-I duplex solution. Then A-I duplex was purified with gel electrophoresis. Different concentrations of thrombin (0-500 nM) and 100 nM purified A-I were mixed and diluted to 100  $\mu$ L in thrombin buffer. Fluorescence was tested after incubation for 3 hours by using Fluorolog (Jobin Yvon Horiba). In the experiment of making calibration curve, different concentrations of FAM-labeled TA-29 (0-100 nM) were prepared in 100  $\mu$ L thrombin buffer solution, and the fluorescence was measured after 3 hours.

### **Validation of Signal Transduction by Fluorescence**

When the signal transduction process in the Inhibitor Generator module was tested, O (100 nM), G (150 nM) and F (200 nM) were mixed in 1 $\times$  TAE-Mg buffer to the total volume of 100  $\mu$ L, and the fluorescence was monitored in the absence and presence of 100 nM DNA-input. The fluorescence intensities of the system with different concentrations of DNA-input under the same condition were also measured. When the signal transduction process in the Threshold Controller module was tested, 40 nM T was added to the solution described above, and fluorescence was monitored with the concentrations of DNA-input at 0, 40 nM, 50 nM, 60 nM, 100 nM and 200 nM. When thrombin was used as input, the buffer was changed to thrombin buffer. Different concentrations of thrombin (0- 500 nM) were first incubated with A-I for 1 hour. Then O (100 nM), G (150 nM), F (200 nM) and T (20 nM or 35 nM) and thrombin buffer were mixed to the total volume of 100  $\mu$ L. Fluorescence intensities were tested after incubation for 3 hours at room temperature.

## Thrombin Catalytic Activity Assay

The hydrolysis experiment of  $\beta$ -Ala-Gly-Arg-p-nitroanilide diacetate (Sigma-Aldrich) was carried out in thrombin buffer. First, 1  $\mu$ L thrombin was taken from 10  $\mu$ M stock solution, diluted to 196  $\mu$ L, and incubated at room temperature for 10 minutes (50 nM). Then 4  $\mu$ L of 0.5 mM chromogenic peptide substrate  $\beta$ -Ala-Gly-Arg-p-nitroanilide diacetate) was mixed with the thrombin solution to make the final concentration of 10  $\mu$ M. The hydrolysis rate of the substrate was determined by monitoring the absorbance at 405 nm by using Cary 100 spectrometry (Varian) with 200  $\mu$ L macro cuvette. The rates for different thrombin concentrations were tested in the same way (Figure 4-3). When testing the inhibition effect of TA-15 to thrombin, thrombin was incubated with different concentrations of TA-15 and TA-29 in buffer for 1 hour. Then chromogenic peptide substrate was added, and the absorbance was monitored at 405 nm (Fig. 4-4). The hydrolysis rate of thrombin with the aptamer circuit was determined in the thrombin buffer containing 1 $\times$ TAE with 100 mM NaCl, 12.5 mM MgCl<sub>2</sub>, 10 mM KCl, and 1 mM CaCl<sub>2</sub>. At the threshold of 100 nM (thrombin concentration), a 20  $\mu$ L sample of 1  $\mu$ M A-I was incubated with 2  $\mu$ L of 10  $\mu$ M thrombin in thrombin buffer for 3 hours. Then, 20  $\mu$ L of 8  $\mu$ M A, G and F, 20  $\mu$ L of 200 nM T, and 20  $\mu$ L of 1  $\mu$ M O were added, and the total volume of solution was made to 196  $\mu$ L with thrombin buffer. The mixture was incubated at room temperature for 6 hours. At the threshold of 200 nM (thrombin concentration), the volume of T was changed to 35  $\mu$ L, and all other conditions were kept constant. Then 4  $\mu$ L of 0.5 mM chromogenic peptide substrate (Final concentration: 10  $\mu$ M) was added, and the hydrolysis rate was determined by monitoring the absorbance of

mixtures at 405 nm at room temperature. The rates for different thrombin concentrations were tested in a similar way (Fig. 4-5).

## Results and Discussion

### DNA Aptamer Circuit Design

To design the DNA aptamer circuit, the sequences of two anti-thrombin aptamers, TA-15 and TA-29, were used as the core sequences. All DNA strands in the circuit consist of recognition domains (13-17 nt) and toehold domains (5-10 nt), and these domains are functionally independent.

The toehold domains were used to initiate the subsequent branch migration reactions. Normally, a 5-nt toehold can reach a displacement rate of  $10^6 \text{ M}^{-1}\text{s}^{-1}$ , which is fast enough for the reactions in the entire system. There are two independent 5-nt toehold domains in this circuit, termed  $b^*$  and  $t^*$ . Toehold  $b^*$  was designed based on the sequence of TA-29, and toehold  $t^*$  is a universal sequence for all other displacement reactions. To differentiate the displacement reaction rates in the Threshold Controller and the Inhibitor Generator modules, a 10-nt long toehold ( $a^*b^*$ ) was used for T.

In the Input convertor module, the A-I duplex was designed based on the sequence of anti-thrombin aptamer TA-29. To effectively cover the 5-nt toehold recognition domain  $b$  and rapidly sense thrombin, two T bases were added to the TA-29 aptamer sequence, and the complementary base pair number between TA-29 and DNA-Input is 12. Based on the fluorescence kinetics results (Fig. 4-6), rapid fluorescence restorations were found with the addition of 200nM and 500 nM thrombin, demonstrating that A-I duplex can be efficiently dehybridized by thrombin.

The recognition domains for branch migration in the O and T duplexes consist of 15-nt ssDNA sequences which are rich in A and T bases. To avoid undesired crosstalk,

a “clamp” was added to the recognition domains, as described by Qian et al.<sup>51</sup> The recognition domain in G was designed with a 13-nt ssDNA which was 2-nt shorter than TA-15. This 13-nt ssDNA does not bind with thrombin, but it can effectively displace TA-15 from the G duplex, allowing the circuit to perform its function normally.

The basic principle of designing Fuel and Fuel' followed the specifications reported by Qian et al.<sup>51</sup> The sequences of Fuel and Fuel' should be designed based on the sequences of the O and A duplexes. The detailed mechanism of the entropy-driven reactions triggered by Fuel is described in Figure 4-2.

### **Validation of Signal Transduction in Each Module**

To ensure proper operation of the entire circuit, the signal transduction in each module was validated separately. The function of the Input Converter is to transform the protein signal to a DNA signal for compatibility. Unlike DNA hybridization, the quantitative binding relationship between protein and aptamer depends on the  $K_d$  of the aptamer and cannot be simply regarded as a 1:1 ratio. Thus, a quantitative relationship needs to be established in order to set an appropriate threshold value for subsequent reactions. Motivated by the design of FRET-based aptamer biosensors,<sup>85, 132</sup> a fluorophore (FAM) in the TA-29 strand and a quencher (DABCYL) in the DNA-input strand were coupled to quantify the released DNA-input triggered by thrombin (Figure 4-7A). The fluorescence enhancement generated by the dehybridization of duplex A-I was measured after addition of different concentrations of thrombin (Figure 4-7B). The resulting calibration curve (Figure 4-8) established the quantitative relationship between thrombin and released DNA-input, as shown in Figure 4-7C. This allowed the quantity of

DNA-input generated from thrombin for subsequent reactions to be determined precisely.

Next, the performance of the Inhibitor Generator was verified by labeling G duplex with a fluorophore and quencher FRET-pair. As a result, when TA-15 is released from G, the fluorescence intensity of the system is restored. For this modular test, DNA-input was directly applied as input to activate the Inhibitor Generator. Before testing, all duplex components were purified using gel electrophoresis to remove excess strands in order to avoid undesired system leakage. Then, O (100 nM), G (150 nM) and F (200 nM) were mixed in the buffer, and the fluorescence was monitored in the absence and presence of 100 nM DNA-input, as shown in Figure 4-9A. Without DNA-input, the Inhibitor Generator is stable for hours. Upon the introduction of DNA-input, an obvious fluorescence enhancement can be observed, indicating that the Inhibitor Generator module functions correctly and effectively. In addition, different concentrations of DNA-input were tested, showing a nonlinear fluorescence restoration with elevated concentration of DNA-input (Figure 4-10). The results demonstrate that the DNA cascade reactions can be driven forward with the catalytic effect of Fuel, thus generating a large amount of TA-15 with a low concentration of DNA-input.

By integrating the Threshold Controller with the Inhibitor Generator, a sequential reaction order with a sharp threshold value was expected. A threshold concentration of 40 nM was set in the Threshold Controller by adding 40 nM purified T duplex. The FRET-based G in the Inhibitor Generator was still applied as a reporter to visualize the signal readout. A series of different concentrations of DNA-input were introduced to verify the efficacy of threshold suppression in the presence of both modules. The

fluorescence kinetics experiments (Figure 4-9B) exhibited fluorescence restoration starting from DNA-input concentration of 50 nM, and the input-versus-fluorescence plot (Figure 4-9C) clearly revealed a sharp threshold value at 40 nM, demonstrating the successful construction of a molecular circuit with precise threshold control.

### **Performance of Entire Circuit with Thrombin as Input**

After confirming the proper function of each component, we further tested the performance of the entire circuit with thrombin as input by fluorescence readout from the FRET-based G. According to the previous DNA-input quantitative results (Figure 4-7C), the concentration of T was set at either 20 nM or 35 nM, corresponding to the concentrations of DNA-input generated by 100 nM and 200 nM thrombin, respectively. When the concentration of thrombin is below the threshold value, only the Threshold Controller is functioning with no fluorescence enhancement. However, once the thrombin concentration exceeds the threshold value, depletion of the Threshold Controller and activation of the Inhibitor Generator occur, thus increasing the fluorescence intensity. In Figure 4-11, the fluorescence restoration started at thrombin concentrations of 100 nM and 200 nM for the circuit with two different threshold concentrations, respectively, confirming that the circuit can function as designed and release TA-15 only when the thrombin concentration exceeds the predetermined threshold value.

### **Programmable and Autonomous Protein Regulation by Aptamer Circuit**

The concentrations of free thrombin observed *in situ* range from less than 1 nM to greater than 100-500 nM.<sup>133, 134</sup> The local concentration of thrombin can even exceed 500 nM with the inducement of some thrombin-generating drugs. These high levels of thrombin cannot be removed from the body in a short time, thus leading to severe blood

clotting. Therefore, a predetermined threshold value, above which the thrombin inhibition drug works, can effectively avoid excessively high thrombin activity, while, at the same time, maintaining regular coagulation function. To demonstrate that our circuit can realize such manipulation of thrombin function, a commercially available chromogenic peptide,  $\beta$ -Ala-Gly-Arg-*p*-nitroanilide diacetate, was chosen as the thrombin substrate. The catalytic activity of thrombin can be determined from the hydrolysis rate of the substrate by monitoring the absorbance of hydrolysis product *p*-nitroanilide at 405 nm.<sup>135-137</sup> The absorbance change ( $\delta A$ ) after 30 minutes was recorded as the evaluation index of thrombin catalytic activity. The influence of the two aptamers, TA-15 and TA-29, on the catalytic function of thrombin was first investigated. The results (Fig. 4-5 and 4-12) confirmed that TA-15 can inhibit the coagulation function of thrombin, while TA-29 has no inhibitory effect on thrombin.

As a result of sub- $\mu$ M binding affinity of TA-15 to thrombin (Figure 4-12), at least 2-fold excess TA-15 is required to obtain more than 50% inhibition effect. In order to generate sufficient signal molecules for downstream reactions, the concentration of O may need to be elevated. However, in this scenario, O would be present at much higher concentration than T, possibly leading to failure of the Threshold Controller caused by the increased displacement reaction rate in the Inhibition Generator. Our initial attempt to adapt the strategy of Figure 4-1 exhibited continuous inhibitory effect on thrombin catalytic activity without any threshold control (Fig. 4-13). To solve this problem, a modified Inhibition Generator module was designed. Because of the modularity of the aptamer circuit, this modified Inhibition Generator module could simply replace the previous Inhibition Generator module without affecting the other two functional modules.

As shown in Figure 4-14, a duplex component, termed Amplifier (A) with its own catalytic strand Fuel', was engineered and incorporated into the module to delay the amplification process for one step. In this case, O can be kept at its regular concentration to generate ssDNA cth in the presence of DNA-input. Then, the ssDNA cth displaces A with the help of Fuel' through an entropy-driven amplification cycle and releases a large amount of htd' strand, thus generating sufficient TA-15 for subsequent inhibition. Since A has no sequence similarity with T, it can be present at high concentration without influencing the effective signal transduction in the Threshold Controller. For instance, the concentration of A in the following regulation reactions was set at 800 nM, which is 20 times higher than that of T. Thus, the strategy of the amplification delay step is to maintain the correct reaction order between the Threshold Controller and the Inhibition Generator, while producing sufficient TA-15 to perform the inhibition function.

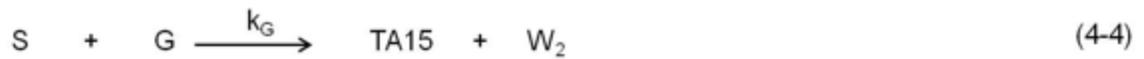
To demonstrate the function of the circuit with threshold control of blood coagulation in the presence of excessively high thrombin concentrations, two parallel threshold concentrations of 100 nM and 200 nM were tested. Below the threshold value, the system should perform the normal coagulation function of thrombin alone. Once the thrombin concentration exceeds the threshold value, inhibition is actuated with the release of TA-15, and the hydrolysis rate of the substrate is attenuated (monitored by absorbance at 405 nm, as shown in Fig. 4-5). The absorbance changes ( $\delta A$ ) after 30 minutes for different thrombin concentrations are summarized in Figure 4-15A. Without the aptamer circuit, larger  $\delta A$  values were found with higher thrombin concentrations. In the presence of the threshold control circuit, but below the threshold concentration,  $\delta A$

kept increasing with increasing thrombin concentration, showing similar behavior to that of thrombin alone. However, once the threshold value had been exceeded, the activity of redundant thrombin was effectively inhibited, resulting in a nearly constant, or slightly increased,  $\delta A$  value compared to that of threshold thrombin. The sharp turning point can be observed exactly at our preset threshold value. Using the 200 nM threshold as an example (green line in Fig. 15A), the green line is nearly identical to the purple line of thrombin alone below 200 nM thrombin input, indicating that the thrombin works freely at this stage. However, in the absence of the circuit, when the thrombin concentration was increased from 200 nM to 500 nM, the  $\delta A$  value was also increased by 141.1%. Otherwise, in the presence of the molecular circuit, it rose only slightly (28.6%), exhibiting a 5-fold difference in activity for the excess portion. A similar trend was observed for a threshold of 100 nM, although the inhibition capability was slightly weaker at higher thrombin concentrations. This is probably caused by the greater total concentration of thrombin needed for inhibition with a lower threshold value, as well as the limited amount of inhibitor the circuit can generate. This problem could be circumvented by presetting with a higher concentration of the Inhibitor Generator module. In addition, we further investigated the initial reaction rates ( $V_{obs}$ ) of this enzymatic reaction. The  $V_{obs}$  calculated from the slope of the initial linear portion of the absorbance kinetics also exhibited consistent results (Fig. 4-15B). These data strongly demonstrated that effective inhibition starts when the threshold is exceeded and only acts on redundant enzymes, confirming the smart inhibitory function of this molecular circuit.

## Theoretical calculations of circuit performance

For rational engineering of an effective molecular circuit, it is important to derive kinetic reaction equations and competitive equilibrium expressions to better understand and predict system behavior. Therefore, we analyzed the reactions in our circuit, which are summarized below. The designed reactions include Eq. 4-1, reversible binding reactions between thrombin and A-I; Eq. 4-2, irreversible thresholding strand displacement reactions with fast  $k_T$  associated with 11-nucleotide extended toeholds; Eq. 4-3, reversible outputting strand displacement reactions with slow forward and backward  $k_O$  associated with 5-nucleotide toeholds, which can be driven forward by a Fuel strand; Eq. 4-4, irreversible inhibitor generation strand displacement reactions with forward  $k_G$  associated with 5-nucleotide toeholds; and, finally, Eq. 4-5, inhibition reactions between TA15 and thrombin. Herein, Eq. 4-1 and Eq. 4-5 are individual reactions which do not influence each other. The elementary steps may be written as:

Based on the theoretical calculations shown in the supporting information, as an independent reaction, Eq. 4-1 can reach the equilibrium within 30 minutes (Fig. 4-16A), which correlates with the practical experiment results (Fig. 4-6), indicating that reaction 1 occurs relatively rapidly. Reaction 4-5 can reach equilibrium even more rapidly than reaction 1, with  $t_{1/2}$  of 250 s based on our simple kinetic predictions (Fig. 4-17B). For Eqs. 4-2, 4-3 and 4-4, as studied by Qian and Winfree,<sup>51</sup>  $k_T = 2 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$  with



the toehold length of 11 nt;  $k_O \approx k_G \approx 5 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$  with the toehold length of 5 nt. The reaction kinetics can be simulated using the built-in numerical integration algorithms.<sup>138</sup> After the simulation, we found that Eqs. 4-2, 4-3 and 4-4 achieve equilibrium in 4-5 hours, which is much slower than reactions 1 and 5. Thus, Eqs. 4-2, 4-3 and 4-4 based on DNA strand displacement are the rate-limiting steps of the system.

For practical applications of the circuit to control the coagulation function of thrombin, a rapid response to the thrombin concentration fluctuation is necessary. To improve system response, it is necessary to accelerate Eqs. 4-2, 4-3 and 4-4. Optimized sequence design may be considered, such as increasing the toehold numbers of O and G in Eqs. 4-3 and 4-4, respectively. But this needs to be balanced with the sensitivity of threshold control. A better solution would be to assemble the circuit modules on a scaffold, where DNA species can interact without diffusion, and the local concentration of DNA strands can be increased.

In addition, the relatively insensitive response of Eq. 4-1 to thrombin decreased the effectiveness of the threshold control in the entire circuit. Further mathematical calculations were performed and revealed that the intrinsic limitation for DNA-input generation is the equilibrium constant  $K_{eq}$  of Eq. 4-1.

To address an important design question regarding obtaining sufficient DNA-input and promoting the high  $\Delta[\text{DNA-input}]$  generation upon different concentrations of thrombin for effective threshold control, we can plot the  $\Delta[\text{DNA-input}]$  triggered by high concentration of thrombin (with  $a=4$ , for example, where  $a$  is the ratio of  $[\text{Tmb}]$  and  $[\text{A-I}]$ ) and low concentration of thrombin ( $a=0.5$ ) with varying  $K_{eq}$  (Fig. 4-17B and 4-17C). Closer examination of the  $\Delta[\text{DNA-input}]$  with varying  $K_{eq}$  reveals a clearly defined local maximum ( $K_{eq} \approx 6.7$ ) that establishes the  $\Delta[\text{DNA-input}]$  values for which the difference between the high thrombin concentration and low thrombin concentration is optimal. To further lower the threshold control for thrombin concentration, a reaction with higher  $K_{eq}$  is desired, and this is related to the binding affinity of the aptamer to thrombin, as well as the dissociation of the A-I duplex. Therefore, a better aptamer with higher affinity to thrombin would lower the threshold control concentration in our circuit. Another possible solution is to reduce the base pair numbers in the A-I duplex for easier release of DNA-input upon target-aptamer binding. But the functional domain on DNA-input still needs to be effectively blocked in the A-I duplex to avoid undesired leaky reactions.

To study the reaction kinetics of these reactions, we have set the original concentration for Tmb, T, O, G and A-I as 300 nM, 40 nM, 100 nM, 150 nM and 100 nM, respectively. The chemical kinetics rate equation for Eq. 4-1 is:

$$k_1[\text{Tmb}][\text{A - I}] - k_{-1}[\text{Tmb-TA29}][\text{DNA - input}] = \frac{d[\text{DNA - input}]}{dt}. \quad (4-6)$$

At any given moment in this reaction, we know that  $[Tmb-TA29] = [DNA-input]$ ,  $[Tmb] = 300 \text{ nM} - [DNA-input]$  and  $[A-I] = 100 \text{ nM} - [DNA-input]$ . Therefore, Eq. 4-6 can be converted as:

$$k_1(300 - [DNA - input])(100 - [DNA - input]) - k_{-1}[DNA - input]^2 = \frac{d[DNA - input]}{dt} \quad (4-7)$$

Integration of Eq. 4-7 gives the integrated form of the rate equation. The equilibrium constant of Eq. 4-1 was obtained as  $K_{eq} \approx 0.12$  from experimental data (Fig.4-7C). Assuming that  $k_1 \approx 3 \times 10^4$  and  $k_{-1} \approx 2 \times 10^5$ , integration of Eq. 4-7 gives:

$$\frac{313 + 7 \cdot [DNA - input]}{300 - 7 \cdot [DNA - input]} = e^{\frac{t+8.5}{200}} \quad (4-8)$$

It is important to note that Eq. 4-3 is driven forward by entropic factors, because a small amount of DNA-input can catalyze the release of a large amount of S. Therefore, the concentration of S is only dependent on the concentration of O. Because Eq. 4-4 is irreversible and the concentration of G is higher than O, the concentration of TA15 is directly correlated to the amount of O. To simplify the model, we assume O duplex can be completely consumed to generate TA15. Therefore, in this case, we can assume  $[TA15]$  is the same as the starting concentration of 100 nM and that  $[thrombin]$  is 300 nM. If we focus on Eq. 4-5, the chemical kinetics rate equation for Eq. 4-5 is:

$$k_2[Tmb][TA15] - k_{-2}[Tmb - TA15] = -\frac{d[Tmb]}{dt} \quad (4-9)$$

At any given moment in this reaction, we know that  $[TA-15] = [Tmb] - 200 \text{ nM}$  and  $[Tmb-TA15] = 300 \text{ nM} - [Tmb]$ . Therefore, Eq. 4-9 can be rewritten as:

$$k_2[Tmb]([Tmb] - 200) - k_{-2}(300 - [Tmb]) = -\frac{d[Tmb]}{dt} \quad (4-10)$$

Integration of Eq. 4-10 gives the integrated form of the rate equation. Here we

know that  $K_2 = \frac{1}{K_d} = \frac{1}{500nM} = 2 \times 10^6 M^{-1}$ ,<sup>132</sup> and assuming that  $k_2 \approx 8 \times 10^3 M^{-1}s^{-1}$ ,<sup>139</sup> then

$k_{-2} = 4 \times 10^{-4} s^{-1}$ . Substitution of these values into Eq. 4-10 gives the final equation as:

$$[Tmb - TA15] = 300 - \frac{22.6 + 265.3e^{0.007t}}{e^{0.007t}} \quad (4-11)$$

The predicted reaction kinetics of Eq. 4-1 and Eq. 4-5 are plotted in Figure 4-17.

As indicated in the Eq. 4-1, the initial concentrations of A-I and Tmb are set to  $[A-I] = c$ , and  $[Tmb] = ac$  where  $a$  is the ratio of  $[Tmb]$  and  $[A-I]$ . The equilibrium concentration of Tmb-TA29 ( $[Tmb-TA29]_{\infty}$ ) is  $x$ . The value of  $x$  must satisfy the equilibrium constant equation :

$$K_{eq} = \frac{[Tmb-TA29]_{\infty} \cdot [DNA - input]}{[Tmb]_{\infty} \cdot [A - I]_{\infty}} = \frac{x^2}{(ac - x) \cdot (c - x)}, \text{ Where } K_{eq} = \frac{k_1}{k_{-1}}.$$

The exact expression for  $x$  (the equilibrium concentration of Tmb-TA29) as a function of  $K_{eq}$  (reaction equilibrium constant between A-I and Tmb),  $ac$  (initial concentration of Tmb) and  $a$  (concentration ratio of  $[Tmb]$  and  $[A-I]$ ) is given by the following equation. Herein, since the equilibrium concentration of DNA-input ( $[DNA-input]$ ) is equal to the equilibrium concentration of Tmb-TA29. Therefore,

$$[DNA - input] = \frac{(a+1) \cdot cK_{eq} + \sqrt{(a-1)^2 \cdot (cK_{eq})^2 + 4ac \cdot cK_{eq}}}{2(K_{eq} - 1)}. \quad (4-12)$$

Herein we fixed  $[A-I] = c = 100$  nM. Figure 4-18A shows plots of  $[DNA-input]$  vs  $K_{eq}$  for different  $a$  values.

To check the maximum  $\Delta[DNA-input]$  with different  $a$  values, one example was calculated. Here we take  $a=4$  and  $a=0.5$ , respectively. Then

$$\begin{aligned} \Delta[\text{DNA - input}] &= f([\text{DNA - input}]_{a=4}) - f([\text{DNA - input}]_{a=0.5}) \\ &= \frac{350Keq + \sqrt{2500Keq^2 + 20000Keq} - \sqrt{90000Keq^2 + 160000Keq}}{2 \cdot (Keq - 1)} \end{aligned} \quad (4-13)$$

Under this condition, when  $Keq \approx 6.7$ ,  $\Delta[\text{DNA-input}]$  achieves the maximum value, as shown in Figure 4-17B and C.

## Conclusions

Design of smart protein manipulation methods with such capabilities as specific recognition of target proteins, precise and autonomous control of protein function, and effective suppression of hyperactive enzymatic effects is an important step towards personalized and intelligent disease treatment.<sup>140</sup> In this study, we have successfully designed a threshold-control molecular circuit and achieved autonomous, self-sustained and programmable manipulation of the catalytic activity of thrombin. This is the first molecular logic circuit based on direct NA-protein interactions for the manipulation of protein activities. Based on the flexible modularity, each module can work relatively independently, while their coordinated operation drives the entire circuit properly. By simply replacing or adding new modules, other functions may be feasible.

Compared with circuits based totally on base-pairing,<sup>52, 118, 141</sup> the introduction of an aptamer provides a direct molecular bridge linking DNA with proteins and enzymatic reactions, enabling precise sensing of the local enzymatic environment and smart regulation *in situ*. Moreover, as a variety of aptamers are either available or can be obtained through SELEX to bind a broad range of targets with tunable binding ability, the aptamer circuit developed here can be used as a powerful tool for constructing ligand-controlled regulation systems tailored to respond to specific targets in defined

situations. Although, in its current state, the circuit works only in the test tube, it can be combined with advanced DNA nanotechnology for further *in vivo* experiments. For instance, all these modules can be molded onto a DNA scaffold, such as DNA origami,<sup>142, 143</sup> to build a complete system and facilitate the coordinated operation. Given the tunable regulation, design modularity and target specificity, the prototype aptamer circuit demonstrated here has the potential to enhance DNA technology with new insights and will broaden the utility of DNA circuits for applications in biology, biotechnology, and biomedicine.

Table 4-1. Sequences of molecular circuit for spectrofluorimetry studies.

	Sequences (5'-3')
TA-29-FAM	FAM-TTAGTCCGTGGTAGGGCAGGTTGGGGTGACT
DNA-input-DABCYL	CAAAAAAAAAACACAACCACGGACTAAA(DABCYL) AAAAA
TA-29	TTAGTCCGTGGTAGGGCAGGTTGGGGTGACT
DNA-input (cba)	CAAAAAAAAAACACA ACCACGGACTAAAAAAAAA
Threshold1(a*b*c*)	AGTCCGTGGTTGTGTTTTTTTTTTG
Threshold2 (c)	CAAAAAAAAAACACA
Output 1(b*c*t*)	GTGGTTGTGTTTTTTTTTTGAGATG
Output 2 (d'tc)	TTGGTGTGGTTGGCATCTCAAAAAAAAAACACA
Fuel (etc)	ACATATCAATTCA TCTCA AAAAAAAAAACACA TGAGA
Generator 1 (t*f*)	TGCCAACACACCAACC-FAM CTCCCG GTG
Generator2(TA-15-DABCYL)	DABCYL-GGTTGGTGTGGTTGGCA
TA-15	GGTTGGTGTGGTTGGCA

Table 4-2. Sequences of molecular circuit for chromogenic peptide substrate hydrolysis monitoring

	Sequences (5'-3')
TA-29	TTAGTCCGTGGTAGGGCAGGTTGGGGTGACT
DNA-input (cba)	CAAAAAAAAAACACA ACCACGGACTAAAAAAAA
Threshold1(a*b*c*)	AGTCCGTGGTTGTGTTTTTTTTTTG
Threshold2 (c)	CAAAAAAAAAACACA
Output 1(b*c*t*)	GTGGTTGTGTTTTTTTTTTTGAGATG
Output 2 (d'tc)	TTGGTGTGGTTGGCATCTCAAAAAAAAAACACA
Amplifier 1(t*h*t*)	ACATATCAATTCA TCTCA AAAAAAAAAACACA TGAGA
Amplifier 2 (d'th)	TTGGTGTGGTTGGCATCTCAAACAAAACCTCA
Fuel' (rth)	CACTCATCCTTTACATCTCAAACAAAACCTCA
Generator 1 (t*f*)	TGAGATGCCAACCACACCAA
Generator2(TA-15)	GGTTGGTGTGGTTGGCA

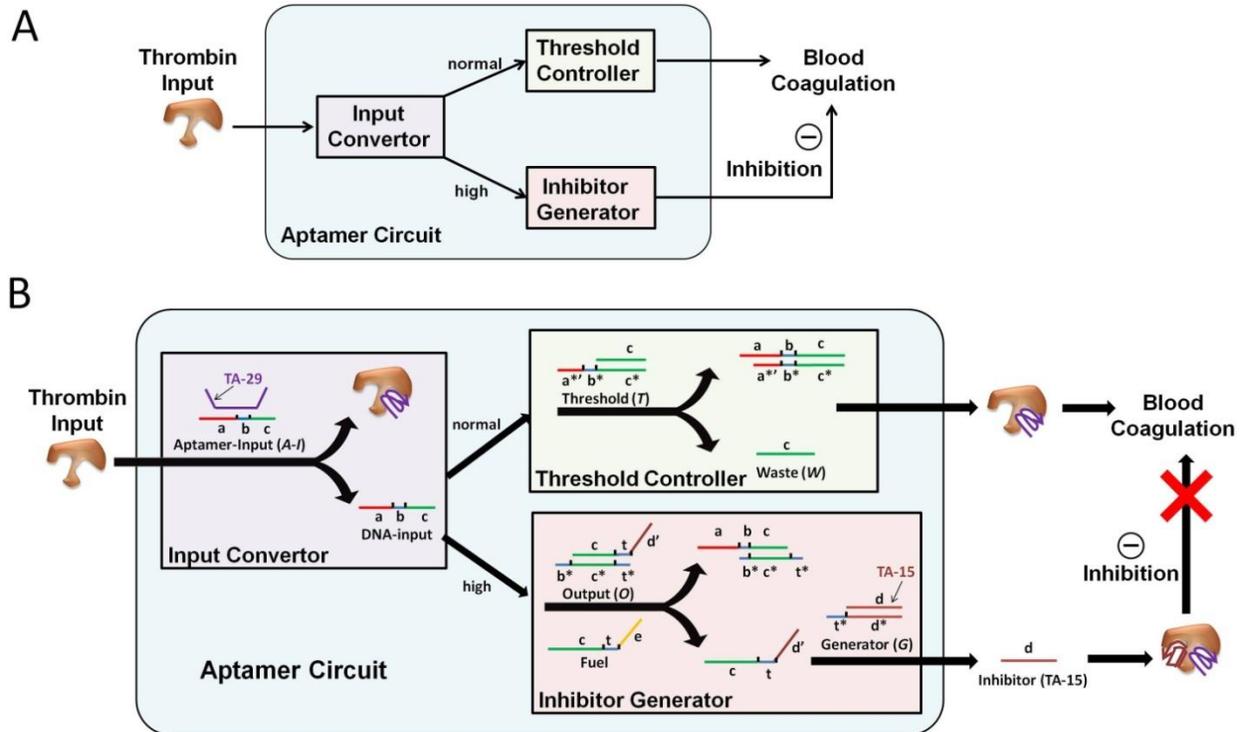


Figure 4-1. Working Scheme of molecular circuit. A) Diagram illustrating circuitry. The circuit consists of three modules, Input Converter, Threshold Controller and Inhibitor Generator, which can be programmed with threshold control for smart manipulation of protein activity. B) Working scheme for molecular circuit, driven by a series of DNA displacement reactions. Colored lines indicate DNA strands with different domains. TA-29 and TA-15 are two thrombin aptamers with different functions, including recognition and inhibition, respectively. All x domains are complementary to  $x^*$ ;  $b^*$  and  $t^*$  are short toehold domains with 5-nt;  $a^*b^*$  is a long toehold domain with 10-nt;  $c^*$  and  $d^*$  are recognition domains with 15-nt. A-I, T, O and G are initially present as duplex components, along with ssDNA Fuel.

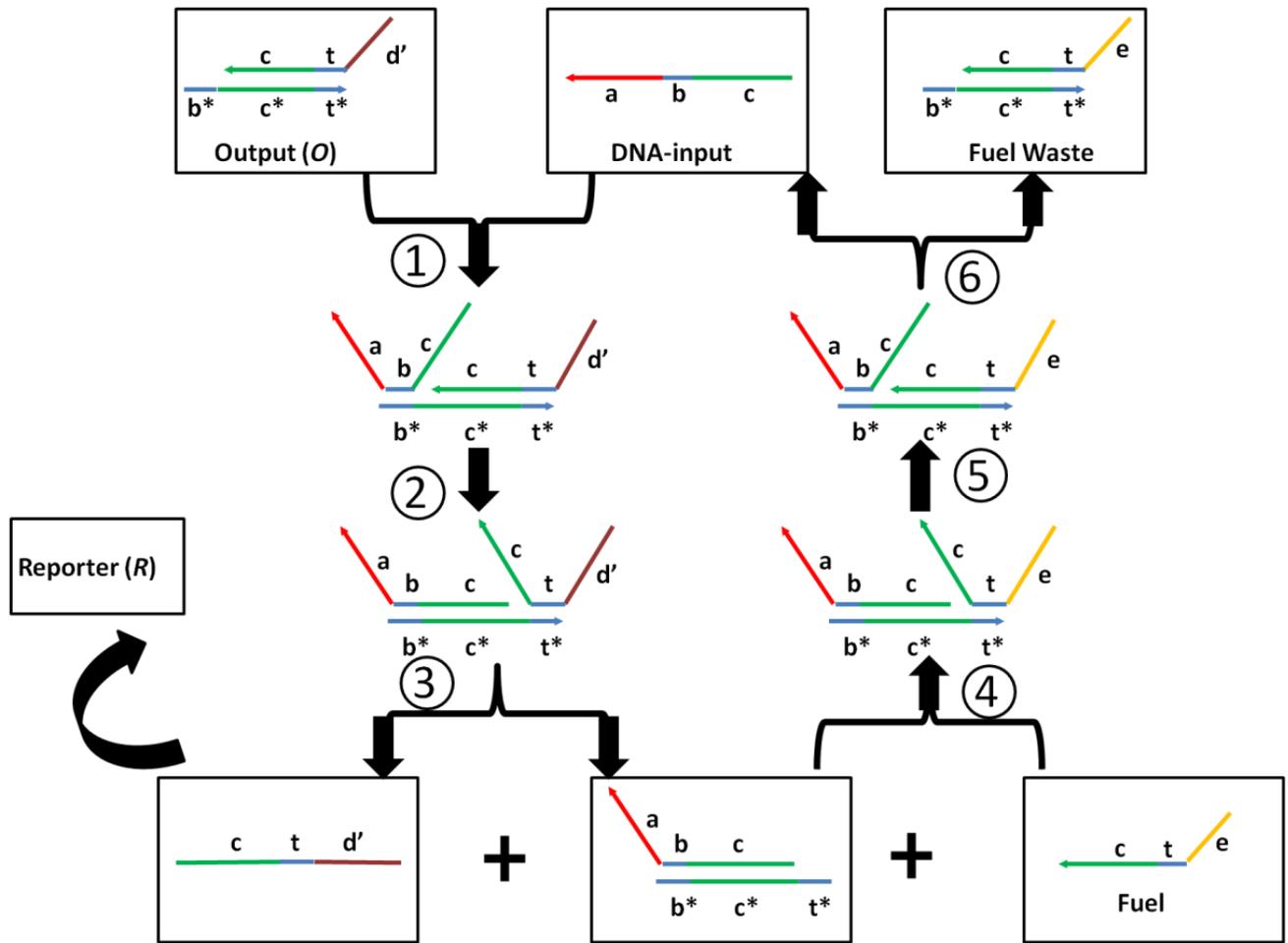


Figure 4-2. Reaction pathways of DNA-input and Output (O) with the aid of Fuel (F). The design follows the procedure described by Qian and Winfree.<sup>51</sup>

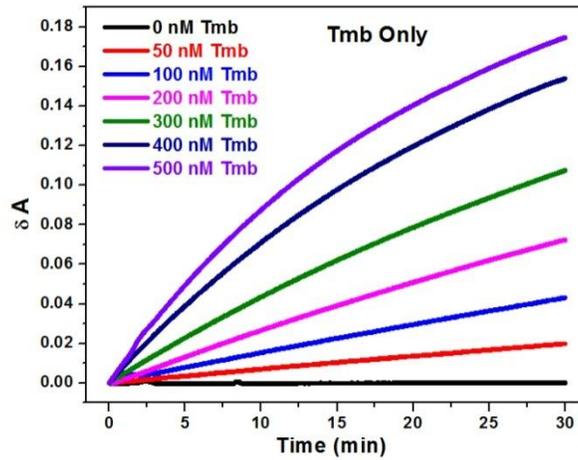


Figure 4-3. Absorbance of chromogenic peptide hydrolysis product vs. time for different concentrations of thrombin.  $C_{\text{peptide}}=10 \mu\text{M}$ ,  $\lambda=405 \text{ nm}$ .

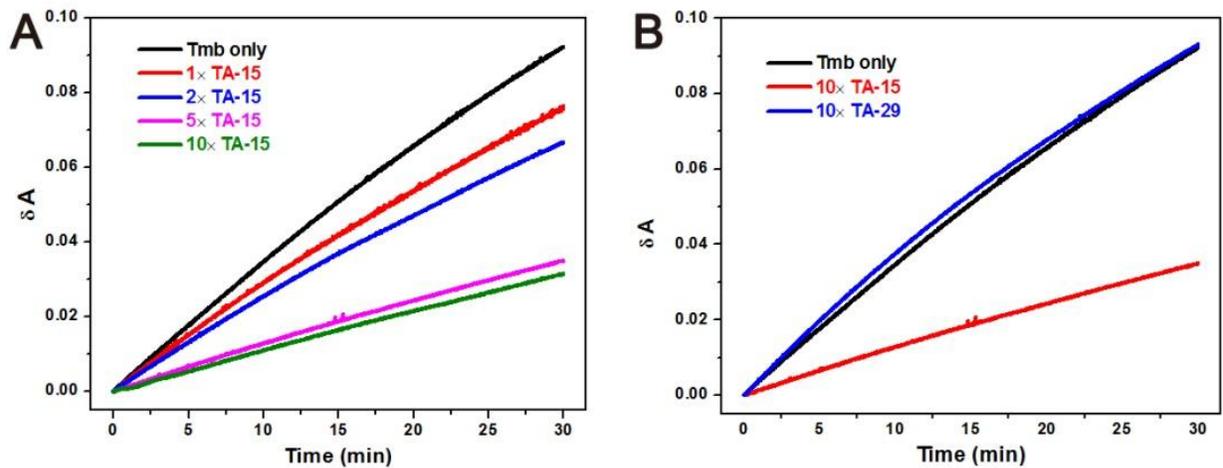


Figure 4-4. A) Absorbance change of chromogenic hydrolysis product as a function of time with 300 nM thrombin preincubated with 150 nM, 300 nM, 600 nM and 1.5  $\mu\text{M}$  TA-15. B) Absorbance change of chromogenic hydrolysis product as a function of time with 300 nM thrombin preincubated with 1.5  $\mu\text{M}$  TA-15 and TA-29.

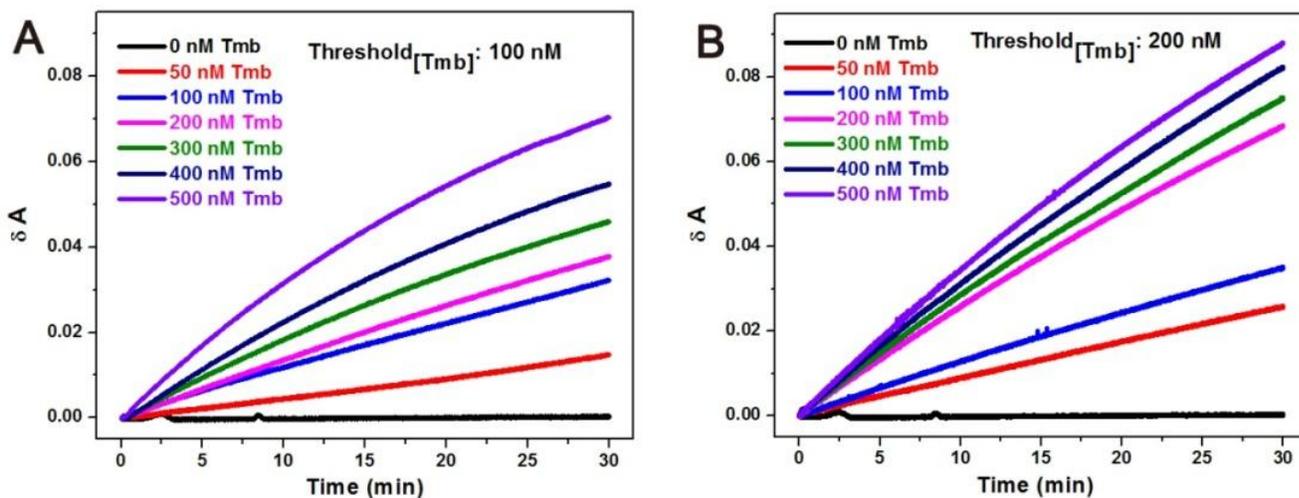


Figure 4-5. Absorbance of chromogenic peptide hydrolysis product as a function of time in the presence of different thrombin concentrations with the aptamer circuit. A) Threshold ( $T_{Tmb}$ ) = 100 nM. B) Threshold ( $T_{Tmb}$ ) = 200 nM. Here, the Inhibitor Generator 2 module was coupled with the other two modules.

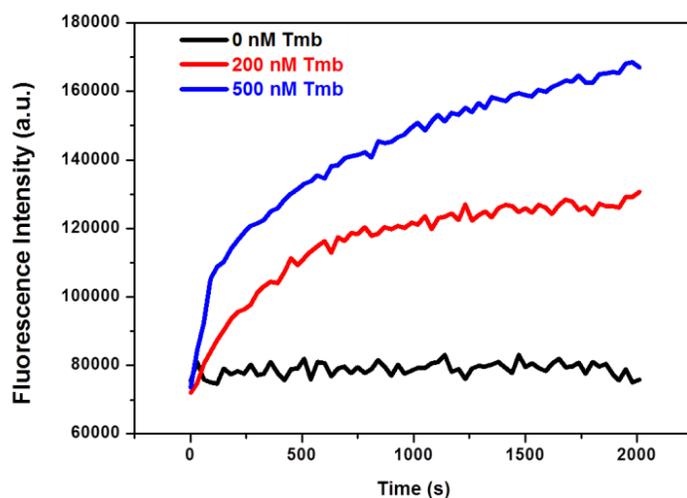


Figure 4-6. Fluorescence kinetics of FRET-based A-I duplex (100 nM) in the presence of 0 nM, 200 nM, and 500 nM thrombin.

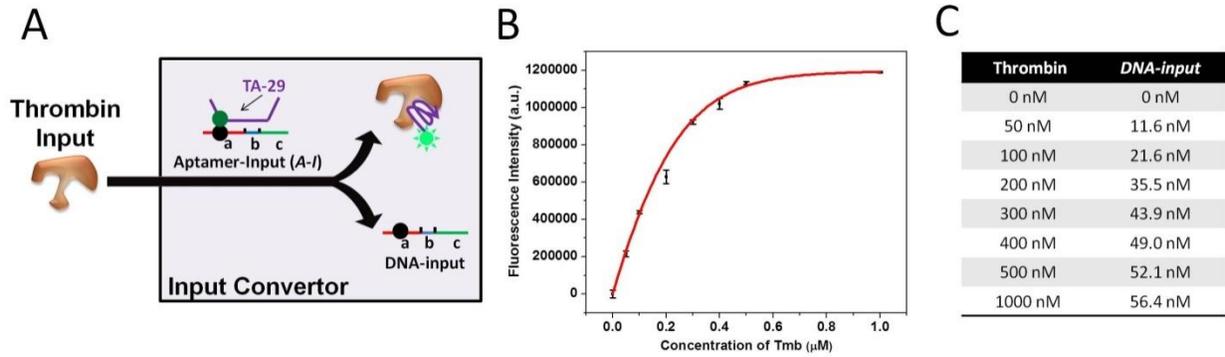


Figure 4-7. Validation of the signal transduction in the Input Converter module. A) Scheme of FRET-based thrombin sensor in the Input Converter module. Quencher-labeled DNA-input and fluorophore-labeled TA-29 can be dehybridized by thrombin. B) Plot of the fluorescence restoration of 100 nM A-I with different concentrations of thrombin. C) Relationship between different concentrations of thrombin and released DNA-input. The concentration of A-I is 100 nM.

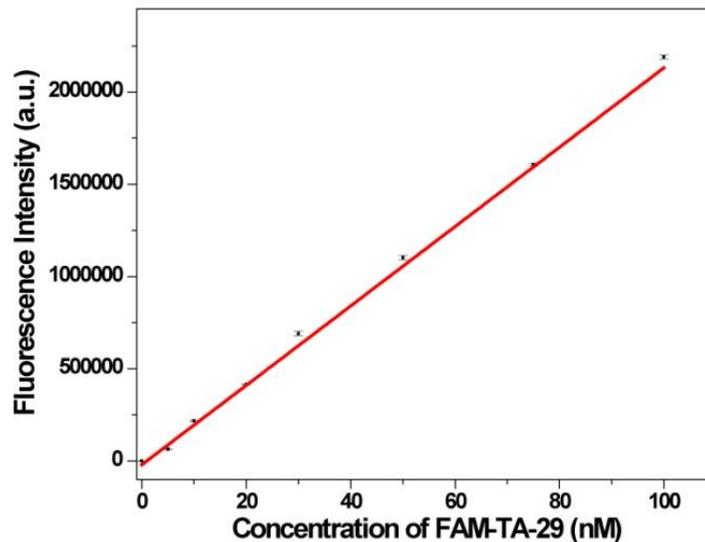


Figure 4-8. Fluorescence calibration curve for different concentrations of FAM-labeled TA-29. The linear range is 0-100 nM. The relationship of fluorescence intensity versus the concentration of FAM-labeled TA-29 was determined as follows:  $\text{Fluorescence} = -19981 + 21511[\text{TA-29}]$ , where  $R^2=0.9877$ .

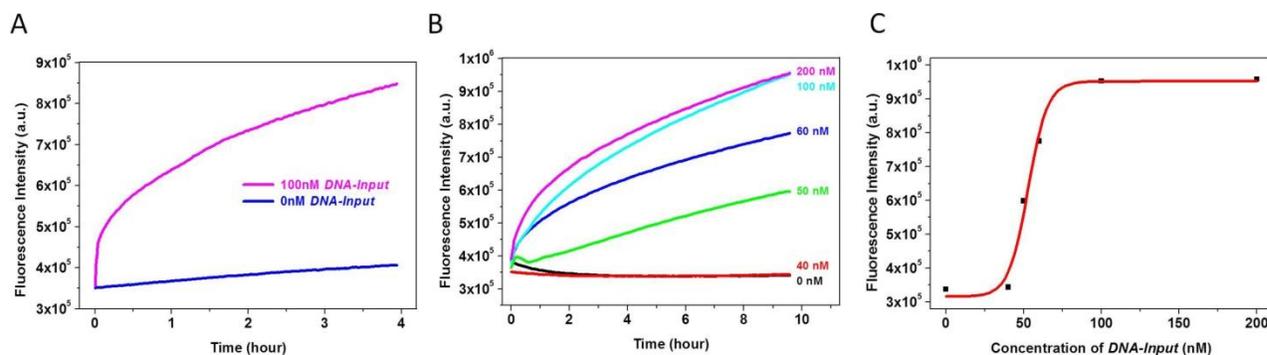


Figure 4-9. Validation of the signal transduction in the Threshold Controller and Inhibitor Generator modules. A) Fluorescence kinetics studies of the signal transduction process in the Inhibitor Generator module. O, G and F were mixed with concentrations of 100 nM, 150 nM and 200 nM, respectively. Then 0 and 100 nM DNA-input were added to the mixture, and the fluorescence signal was monitored at 25°C. B) Fluorescence kinetics study of the signal transduction of the Threshold Controller together with the Inhibitor Generator. T, O, G and F were mixed with concentrations of 40 nM, 100 nM, 150 nM and 200 nM, respectively. ssDNA DNA-input was added at different concentrations (0 nM, 40 nM, 50 nM, 60 nM, 100 nM and 200 nM), and the fluorescence was monitored with  $\lambda_{ex}$ (488 nm) and  $\lambda_{em}$ (518 nm). C) Plot of output final fluorescence intensity versus concentration of DNA-input.

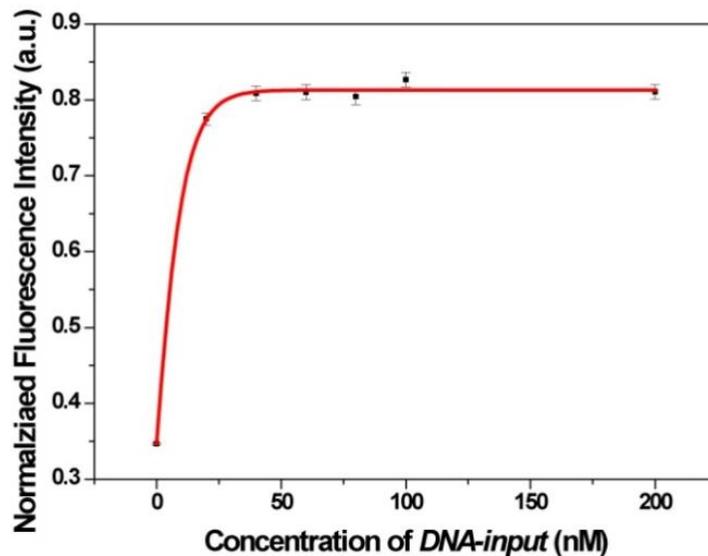


Figure 4-10. Fluorescence output versus DNA-input concentration plot. The final fluorescence was normalized to the maximum completion level. These results indicate that 20 nM of DNA-input can release around 80% of TA-15 from G duplex.

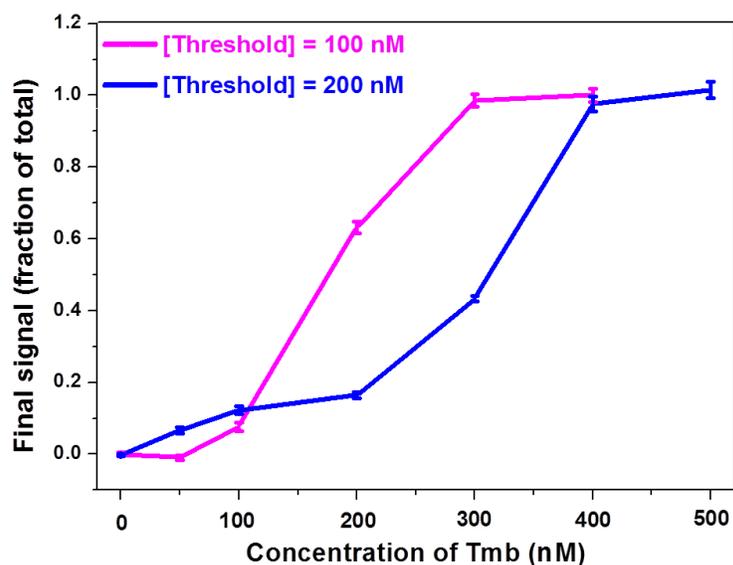


Figure 4-11. Fluorescence output versus thrombin concentration plot with two preset threshold values (concentration of Tmb) at 100 nM and 200 nM. Purified A-I, O, G and F were mixed at the concentrations of 100 nM, 100 nM, 150 nM and 200 nM, respectively. Fluorescence signals were recorded with  $\lambda_{\text{ex}}$ (488 nm) and  $\lambda_{\text{em}}$ (518 nm) at 25°C. The final fluorescence was normalized to the maximum completion level.

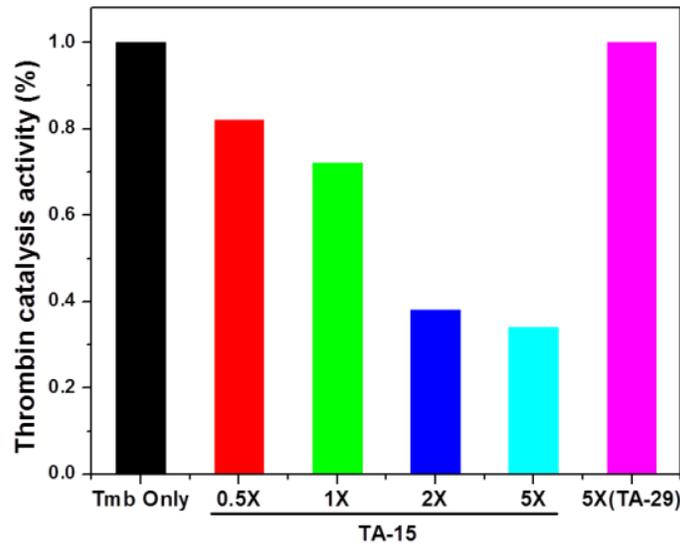


Figure 4-12. Inhibitory effects of different concentrations of TA-15 and TA-29 on thrombin. The concentrations of thrombin and chromogenic peptide substrate are 300 nM and 10  $\mu$ M, respectively.

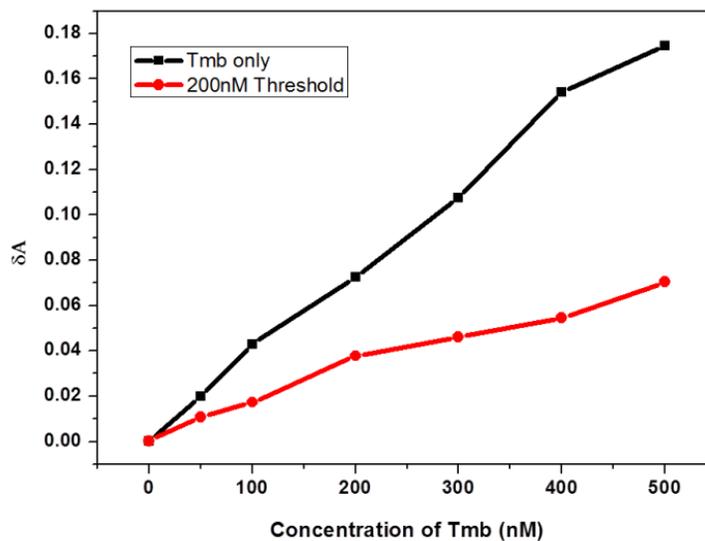


Figure 4-13. Absorbance changes of chromogenic peptide substrate for different thrombin concentrations using the design in Figure 4-1. Here, the Inhibitor Generator module was coupled with the other two modules.

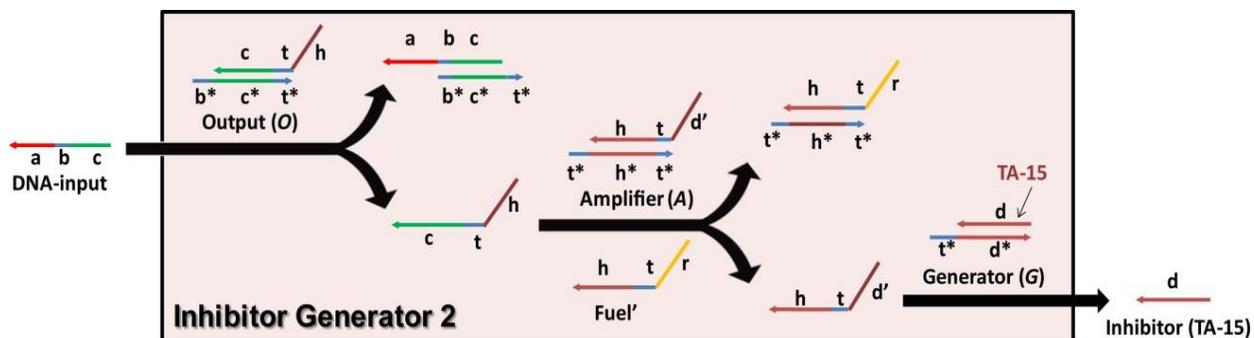


Figure 4-14. Working scheme of Inhibitor Generator 2. Duplex A and the Fuel' strand were designed to delay the amplification process for one step.

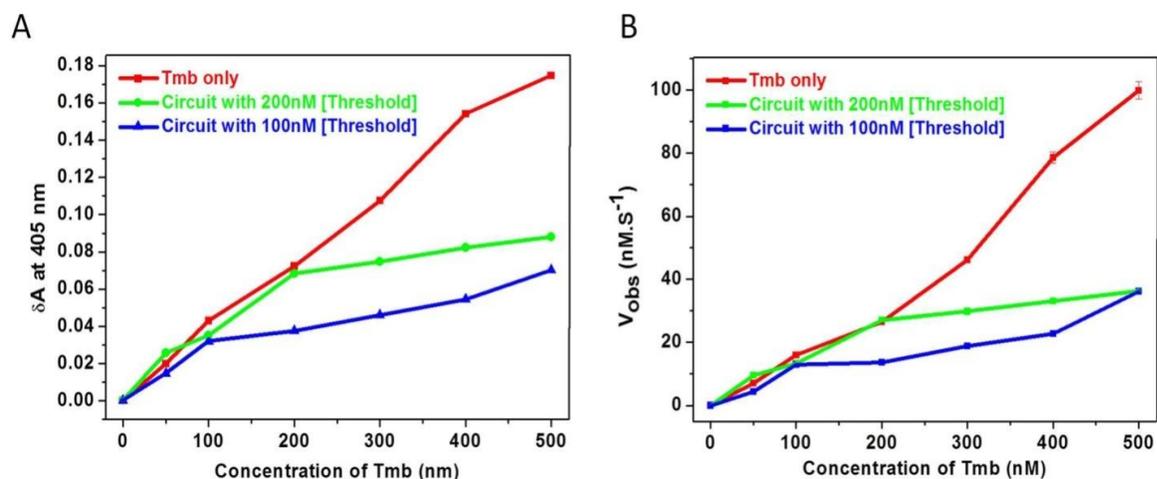


Figure 4-15. A) Absorbance changes ( $\delta A$ ) of chromogenic peptide substrate upon different thrombin concentrations with and without circuit. B)  $V_{obs}$  of the enzymatic reaction upon different thrombin concentrations with and without circuit. Purified A-I, O, A, G and F were mixed at concentrations of 100 nM, 100 nM, 800 nM, 800 nM and 800 nM, respectively. T was set at either 20 nM or 35 nM for circuits with two different threshold values.

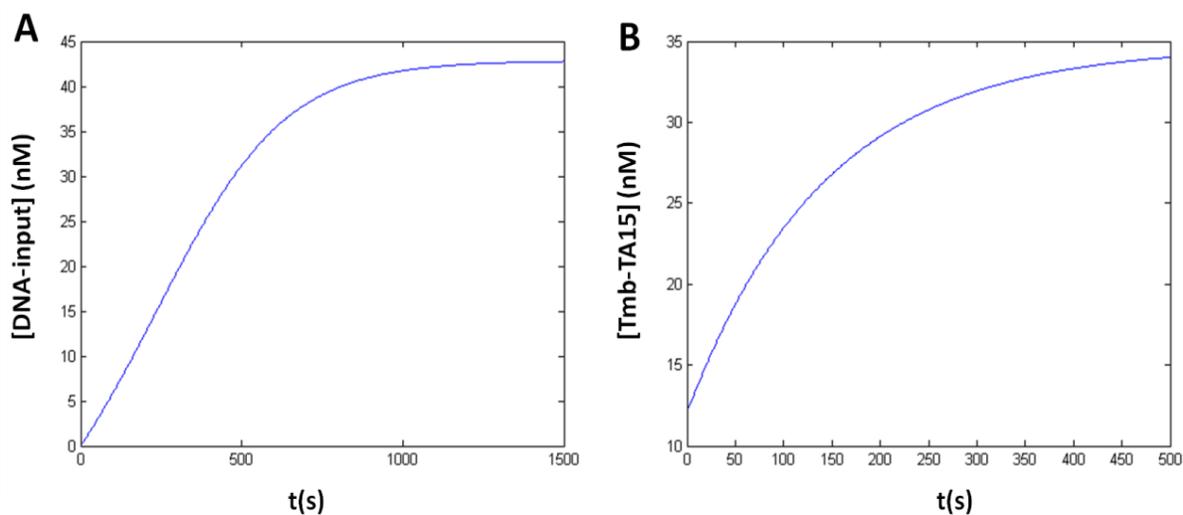


Figure 4-16. A). Plot of theoretical [DNA-input] vs. time. B). Plot of theoretical [Tmb-TA15] vs. time.

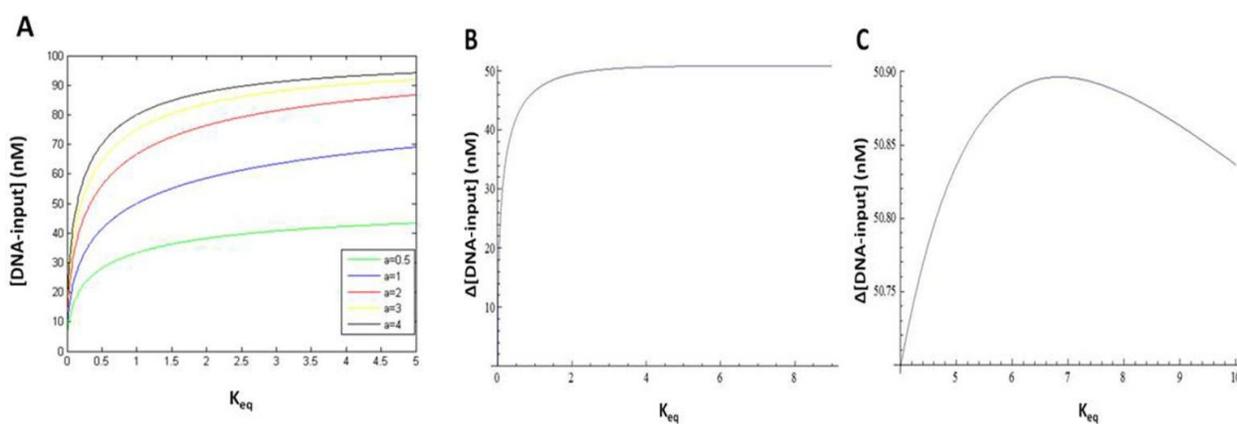


Figure 4-17. A) Plot of [DNA-input] vs.  $K_{eq}$  for different  $a$  values, where  $a$  is the concentration ratio of [Tmb] and [A-I]. B) Plot of  $\Delta$ [DNA-input] vs.  $K_{eq}$ . Here,  $\Delta$ [DNA-input] was calculated by using  $f$  ([DNA-input] <sub>$a=4$</sub> ) –  $f$  ([DNA-input] <sub>$a=0.5$</sub> ). C) The zoom-in plot of (B).

## CHAPTER 5 MOLECULAR ENGINEERING OF A DNA-ENZYME CASCADE NETWORK FOR ACQUIRED IMMUNE SYSTEM MIMICRY

### **Significance and Background**

Acquired, or adaptive, immune system (AIS) is controlled by an intricate network of chemical reactions and cellular communications. Macroscopically, AIS displays behaviors including pattern recognition, antigen tolerance, pathogen elimination and memory formation.<sup>144</sup> Aiming at further understanding their detailed regulatory biological processes which are expected to reveal the underlying design rules through a bottom-up approach, people are looking for suitable hardware to implement the complex reaction networks. Until now, a fully implemented bottom-up AIS biomimicry system still lacks experimental demonstration. Difficulties include controlling programmability and systematicness, as well as transferring the high-level biological codes into simple artificial modules and physical implementations. Another difficulty involves forming robust molecular structures and precisely controlling their temporal dynamics.

As a carrier of genetic information with well-regulated and predictable structures, DNA offers an excellent platform for the design of reaction networks with arbitrary topologies and high programmability.<sup>145, 146</sup> Building on the richness of DNA computing<sup>51, 54</sup> and DNA-enzyme manipulation,<sup>147, 148</sup> synthetic DNA-based systems have been used to explore the possibilities of mimicking both simple and complex systems. For example, small-scale *in vitro* circuits encoding elementary functions, such as cascades<sup>149</sup>, bistable memory<sup>150, 151</sup>, or oscillations<sup>58</sup>, have successfully been engineered. In addition, larger networks including mimicry of the neural network<sup>50</sup> and predator-prey ecosystem<sup>59, 60</sup> continued to confirm the excellence of DNA biochemistry for biomimicry implementation. To further expand the engineering capability of DNA

molecules for building a complex biological network, we rationally designed a simple, but effective, experimental approach for the biomimicry of the vertebrate Acquired Immune System (AIS) and analyzed its robust dynamic behaviors. Our results suggest that DNA strand displacement cascades,<sup>152</sup> coupled with DNA-enzyme interactions,<sup>153</sup> could be used as components to build a general strategy for mimicking complex biological systems, thus driving the future development of biomimicry.

## **Experimental Materials and Methods**

### **DNA Synthesis**

The DNA sequences were synthesized on the ABI 3400 DNA synthesizer. The synthesis protocol was set up according to the requirements specified by the reagents' manufacturers. Following on machine synthesis, the DNA products were deprotected and cleaved from CPG by incubating with 2.5 mL AMA (ammonium hydroxide/Methylamine 50:50) for 17 hours at 40°C in water bath. The cleaved DNA product was transferred into a 15 mL centrifuge tube and mixed with 250 µL 3.0 M NaCl and 5.0 mL ethanol, after which the sample was placed into a freezer at -20 °C for ethanol precipitation. Afterwards, the DNA product was spun at 4000 rpm under 3°C for 20 minutes. The supernatant was removed, and the precipitated DNA product was dissolved in 500 µL 0.2 M triethylamine acetate (TEAA Glen Research Corp.) for HPLC purification. The HPLC purification was performed with a cleaned Alltech C18 column on a Varian Prostar HPLC machine. The collected DNA product was dried and processed detritylation by dissolved and incubated in 200 µL 80% acetic acid for 20 minutes. The detritylation DNA product was mixed with 400 µL ethanol and dried by a vacuum dryer. The DNA products were quantified and stored in DNA water for

subsequent experiments. The detailed sequences information are described in supporting information (Table 5-1).

### **DNA Purification**

Native PAGE was applied to purify the AM, BM and PG duplex strands to remove excess strands and avoid undesired system leakage. The ssDNA components of AM, BM and PG were annealed at concentrations of around 50  $\mu\text{M}$  in 1 $\times$ TAE-Mg buffer (20 mM Tris-Acetate-EDTA, pH 7.5, 12.5 mM  $\text{Mg}(\text{Ac})_2$ ). Native PAGE gels (10%) in 1 $\times$ TAE-Mg buffer were run at 100 V for 90 minutes at 4°C and stained with Gel Red stain solution (Biotium, CA). Only the sharp bands were cut from the gels, chopped into small pieces, and soaked in 1 $\times$ TAE-Mg buffer for 24 hours. After soaking out most DNA molecules from the gel pieces, the solutions were extracted and concentrated with centrifugal filter devices (Millipore, MA). Finally, the DNA duplex sequences were quantified by UV spectrometry and kept in buffer for future use.

### **Preparation of Circular Template DNA (CP)**

A 7.2  $\mu\text{L}$  sample of 20  $\mu\text{M}$  CP (with phosphate group on its 5 prime end ) was placed in 4  $\mu\text{L}$  CirLigase II 10X Reaction Buffer containing 330 mM Tris-acetate (pH 7.5), 0.66 M KAc, 5 mM Dithiothreitol (DTT) and 2.5mM  $\text{MnCl}_2$ . Afterwards, 2.8  $\mu\text{L}$  CirLigase II (Epicentre, WI) was mixed with the buffer and diluted to 40  $\mu\text{L}$  with water. The mixture was incubated at 60 °C for 3 hours. Then 20U Exonuclease I and 200 U Exonuclease III (New England Biotech, MA) were added to the mixture and incubated at 37°C for 1 hour. The enzyme was denatured by heating up the solution to 90 °C for 20 minutes. Finally, the product CP was purified by denatured-PAGE and desalted with NAP-5 columns (Fig.5-10).

## **Validation of Signal Transduction in Each step and the Entire System by Fluorescence**

When the reaction priority between the Step 1 and Step 2 was tested, purified AM (80 nM), BM (100 nM) and PG (with FAM and DAB, 100 nM) were mixed in 1×TAE-Mg buffer to the total volume of 100  $\mu$ L, and the fluorescence was monitored in different concentrations of  $P_0$ . (Fig. 5-4) Fluorescence intensities were tested after incubation for 1 hours at room temperature. When the signal transduction in the entire system was tested by fluorescence, purified AM (80 nM), BM (100 nM), PG (No FAM or DAB, 100 nM), 50 nM CP and 500 nM MB-R were incubated in 1×TAE-Mg buffer. The fluorescence kinetics was started when different amounts of  $P_0$  were added to the solution. To check the fluorescence kinetics of the system with the second-time exposure to  $P_0$ , 50 nM  $P_0$  was firstly added to the buffer containing purified AM (100 nM), BM (80 nM), PG (No FAM or DAB, 100 nM) and 50 nM CP. Then another 150 nM  $P_0$  together with 500 nM MB-R were mixed together in the solution simultaneously.

## **Electrophoresis Analysis of the System by Agarose Gel**

A 10  $\mu$ L system with purified AM (200 nM), BM (200 nM), PG (200 nM), CP (50 nM) Phi29, (0.5U/  $\mu$ L) and dNTP (250  $\mu$ M) was chosen for this experiment. The reaction was carried out in 1×RCA buffer (50 mM Tris-HCl, 10 mM  $MgCl_2$ , 10 mM  $(NH_4)_2SO_4$  and 4 mM DTT) with different amount of  $P_0$  input at 30°C for 3 hours. Agarose gel was run at 4°C for 1 hours under the voltage of 100V.

## **Results and Discussion**

### **System Design and Construction**

The host innate immune system is the first line of defense against invading pathogens, but its effect is nonspecific and short-lived. In contrast, the acquired immune

system consists of very specialized cells and processes to eliminate pathogenic colonization. This mainly involves B- and T-cells, but also the production of antibodies and complement system, which helps antibodies and phagocytic cells eliminate invading pathogens from the body. More specifically, the acquired immune system recapitulates three steady states, including “recognition and tolerance”, “immune response” and “killing and memory”. In the first phase, specific “nonself” antigens in the presence of “self” components are sorted out through the process of antigen presentation, mainly with the help of antigen presenting cells (APC). Afterwards, the response signal is transferred to T cells which can be differentiated to some subcategories, including helper T cells (positive) and suppressor T cells (negative). The mutual stimulation of positive and negative T cells controls the tolerance level of adaptive immunity. However, once the threshold is overcome and tolerance is exceeded, the system moves to the second phase. The dominant existence of helper T cells will transfer the signal to B cells, which are tailored to quickly release antibody and eliminate specific pathogens. Apart from immune response of antibody to antigen, the system will next develop a memory effect in the third phase. Here, particular memory B cell will be remained to “remembered” specific pathogen and used to trigger much faster immune response with subsequent exposure to the system.

To mimic such a complex system, we designed a DNA-Enzyme hybrid system consisting of four individual DNA duplex components and two enzymes (Fig. 5-1, species in the shadow box) able to respond to incoming pathogen input (ssDNA or ssRNA) autonomously and programmably. When no pathogen is present, the system is maintained in a steady, balanced state by effective blocking of the functional domains in

each component. However, when challenged by pathogenic input, these functional domains are activated in a series of steps designed to mimic the three phases of adaptive immunity, as described above. From an engineering perspective, the first step is based on strand displacement, which allows a new sequence to be released by some initiator strands, thus exhibiting complex computational and information processing abilities for the construction of specific chemical reaction networks.<sup>74</sup> In our case, a specific ssDNA pathogen input ( $P_0$ ), can be recognized by the DNA duplex AM (stands for APC cell Mimicry) in step 1 or BM (stands for B cell Mimicry) in step 2 through DNA displacement reaction, in which a single-stranded DNA (ssDNA) can be displaced from the initial duplex by an even stronger binder, normally a better matched DNA strand. Specifically,  $P_0$  comprises of two ssDNA domains with one taken from *Bacillus anthracis* genome (P) and another designed (domain 2-3-4) for controlling downstream reactions. The reaction priority of  $P_0$  to AM and BM is controlled through the lengths of their corresponding toeholds. Herein, we designed a 10-nt toehold in AM with the displacement reaction rate  $k$  of  $10^6$  and a 0-nt toehold in BM with  $k$  value of  $10^3$  for initial pathogen binding.<sup>48, 49</sup> As AM is depleted by displacement, one of the displaced products ssDNA TM (stands for T helper cell Mimicry) accumulates as can be used as a catalyst to increase the reaction rate between  $P_0$  and BM in step 2. At the “recognition and tolerance” step, the system is able to proceed to the next step only when pathogen accumulation exceeds the threshold control ability of AM.

In the second step, the “immune response” state is activated. Here,  $P_0$  can displace BM with the help of TM and release ssDNA PI (stands for primer initiator), which then undergoes another step of DNA displacement with duplex PG (stands for

primer generator), producing ssDNA 12a to become, in turn, the primer for DNA polymerase-catalyzed rolling circle amplification (RCA), a well-established isothermal process that can rapidly synthesize multiple copies of desired ssDNA.<sup>154, 155</sup> Here, we incorporated two *Bacillus anthracis* genomic sequence fragments (P) in the circular template (CP) for RCA, resulting in rapid generation of the complementary sequences of pathogen (antibody P\*) in the presence of Phi29 DNA Polymerase, the replicative polymerase from the *Bacillus subtilis* phage phi29, and deoxyribonucleotide triphosphate (dNTP). Using such enzymatic amplification, this system can quickly generate multiple copies of P\* strands that can mimic the fast and specific antibody generation produced by B cells. Because the RCA product is dependent on its primer concentration, we can control the antibody P\* amount by monitoring the upstream reactions that release primer 12a.

In the last step, the active pathogen (domain P) exists in three infectious statuses, including original pathogen P<sub>0</sub> and displacement reaction products P<sub>1</sub> and P<sub>2</sub> (shown in the red dotted rectangular in Fig. 5-1B). With the large amount of antibody P\* in the system, P domain in P<sub>0</sub>, P<sub>1</sub> and P<sub>2</sub> can form stable duplex regions PP\*. Because of the presence of restriction enzyme SspI, extracted from an E. coli strain that carries the cloned and modified (Y98F) SspI gene from *Sphaerotilus* species, the duplex PP\* will be specifically cut into short strands, resulting in the pathogen's loss of infectivity. This step is used to mimic the process by which the adaptive immune response eliminates pathogens after the formation of antibody-antigen complex from the body. In addition, since the ssDNA TM is remained in the system from the first exposure to pathogen, it can catalyze the P<sub>0</sub> and BM displacement to release primer 12a quickly,

thus triggering a faster RCA reaction for the next exposure to the same pathogen sequence. This reaction is based on a mechanism called entropy-driven behavior, in which the entropy change of reaction system is positive (Fig. 5-2). Without TM, the displacement rate between  $P_0$  and BM is slow by the effective blocking of active domains on BM. However, in the presence of TM, BM can bind with TM through its 4-nt toehold (domain 5\*) and then displace ssDNA  $W_1$  (34), resulting in a new toehold binding region (domain 3\*) for the hybridization between  $P_0$  and BM. Under these conditions, the BM and  $P_0$  hybridization reaction is driven forward thermodynamically by the entropic gain of the liberated molecules. This mechanism promises the system to form memory effect for specific pathogen input through leaving TM as memory B cells in the system.

In addition, the specificity of this AIS mimicry system is mainly dependent on the sequence design of circular template CP as well as the choice of restriction enzyme. Here, because of the encoding of *Bacillus anthracis* genomic sequence fragments (P) in CP, the generated RCA product (antibody  $P^*$ ) is only able to recognize pathogen with P, thus providing the first layer of systematic selectivity. The second layer selectivity is added through the specific sequence digestion requirement of the restriction enzyme SsPI. In other words, even if unspecific binding between other pathogen and antibody  $P^*$  happens, the pathogen sequence cannot be cut by the restriction enzyme SsPI. Therefore, with such designs, our AIS mimicry system possesses the character of specific immune response in the real AIS.

### **Validation of Signal Transduction in Each Module**

To ensure proper operation of the entire network, signal transduction at each step was validated separately. First, AM primarily binds with  $P_0$ , whereas delays the

reaction with BM due to the short toehold region on BM. Therefore, a FRET-based method was used to study the reaction order of  $P_0$  with AM and BM. A fluorophore (FAM) and quencher (DABCYL) pair was coupled in the PG duplex to indicate the released amount of 12a strand that served as a primer for RCA reaction. Fluorescence restoration only occurs if the threshold level of AM is exceeded, and the reaction proceeds to step two, i.e., immune response, based on displacement reactions that can release the quencher-labeled primer 12a strand. Therefore, in the presence of 80 nM AM, 100 nM BM and 100nM fluorophore-quencher labeled PG, different concentrations of  $P_0$  (from 0 nM to 150 nM) were introduced to the system with fluorescence monitoring at 517 nm. As shown in Figure 5-3A, the fluorescence restoration exhibited a sharp upturn at the  $P_0$  concentration of 80 nM, indicating that the reaction between  $P_0$  and BM had begun at the depletion point of AM. In other words, the threshold value of AM to  $P_0$  is adjustable by changing the concentration of AM. To further confirm this result, fluorescence kinetics of the above system was monitored (Fig.5-3B). Initially, 80 nM AM, 100 nM BM and 100 nM PG, and a small amount of  $P_0$  (30 nM) was added. Because of the excess amount of AM and its reaction priority to  $P_0$  by the requirements of step 1, slow fluorescence restoration kinetics was discovered. However, when more  $P_0$  (30 nM each time) was continuously added to the system, a sharp increase was seen in the fluorescence kinetics, indicating that the reaction proceeds to step 2. Different threshold values were also tested by changing the concentrations of AM and BM (Fig. 5-4). Gel electrophoresis was also used to indicate the order of reactions (Fig. 5-5). All assays confirmed the effective thresholding function of AM to the reaction

between BM with  $P_0$ , thus providing a macroscopic mimicry of the “recognition and tolerance” state of AIS.

Another significant function of AM is to produce catalytic ssDNA TM to facilitate  $P_0$  to BM hybridization. As demonstrated by Zhang et al.,<sup>19</sup> the addition of catalyzing TM can accelerate the hybridization reaction between  $P_0$  and BM by two to four orders of magnitude. Therefore, to prove the catalytic effect of TM, we continued to use our FRET-based reporting model, as described above. Accordingly, 100 nM BM, 80 nM AM and 100 nM FAM-DAB-labeled PG were incubated in buffer. Afterwards, 150 nM  $P_0$  were added to the system. It should be noted that we used an excess amount of  $P_0$  to AM in order to overcome the threshold value. Thus, excess  $P_0$ , as well as catalyzing TM generated as a result of the reaction between  $P_0$  and AM in step 1, form an entropy-catalytic reaction that can quickly restore the fluorescence of the system. As a comparison, the same concentration of  $P_0$  was directly added to a solution with 100 nM BM and 100 nM FAM-DAB-labeled PG. As shown in Figure 5-3C, measurements of the fluorescence kinetics of the catalyzed reaction exhibited over 500-fold acceleration in contrast to that of the uncatalyzed.<sup>49, 156</sup> Thus, the released TM can serve as memory mimicry in the system to memorize specific pathogen, resulting a stronger and faster immune response for the next exposure.

After confirming the reaction priority between step 1 and 2, we studied the enzymatic amplification process in step 2, host immune response. As an efficient isothermal enzyme-based amplification, rolling circle amplification (RCA) provides us with an excellent way to specifically produce a large amount of desired DNA product in a short time. Therefore, we applied RCA to mimic the generation of antibody in our

network. By encoding the active pathogen sequence P in the circular template (CP) for RCA, DNA polymerase Phi29 could specifically elongate the primer of 12a from its 3'-end. To demonstrate that the concentration of RCA product is dependent on the amount of primer 12a, different concentrations of ssDNA 12a were incubated with 30 nM CP, 0.5 U/ $\mu$ L Phi29 and 250  $\mu$ M dNTP in RCA buffer. To perform this experiment, we designed a DNA molecular beacon (MB-R) to report the amount of RCA product fluorescently by using pathogen *Bacillus anthracis* genomic sequence P as its loop. MB-R can be opened by the RCA products (antibody P\*), thereby exhibiting fluorescence restoration that represents different amounts of antibody P\*. Figure 3a shows the resultant fluorescence curve establishing the quantitative relationship between antibody P\* and released primer 12a. This allowed the quantity of antibody P\* generated from enzymatic amplification for hybridization with domain P in P<sub>0</sub>, P<sub>1</sub> and P<sub>2</sub> to be determined by the upstream reactions (step 1 and step 2).

We wanted to avoid the undesired elongation of Phi 29 at the 3'-end of any duplex in this system. Consequently, we modified the inverted dT group at the 3'-end of each DNA component,<sup>157, 158</sup> except primer 12a, because 12a is a primer for RCA and should not be initially blocked. However, a problem could arise under these circumstances in that the toehold of PG duplex might be sealed when initially incubated with Phi 29. To solve this problem, we designed 1 or 2 mismatch points at the 3'-end of primer 12a. These mismatched PG duplexes labeled by FRET pairs (FAM-DABCYL) were then incubated with 0.1 U/ $\mu$ L or 1 U/ $\mu$ L Phi 29 for 1 hour. Afterwards, different amounts of ssDNA PI (upstream displacement strand) were added to the system to explore fluorescence restoration. If the toehold of PG is sealed by Phi29, results will

show that the displacement reaction between PI and PG are negligible, leading to low fluorescence recovery. As shown in Figure 5-6A, 2 mismatched bases provided the best anti-elongation effect by showing the greatest fluorescence restoration. We did not test 3 or more mismatches in PG because the increasing instability of PG might have caused greater signal leaks in the system. As a tradeoff against too much elongation, but too little stability, we chose 2 mismatch bases in PG duplex for the implementation of our system.

To mimic the final pathogen elimination process of AIS (step 3), we designed a restriction enzyme-based step to specifically cut the “antibody-antigen” complex. Herein, the hybridization of P domain with the P\* domain in antibody P\* creates an SsPI restriction site. In this way, P could be specifically cut into a fragment with the base numbers of 56 (Major band).<sup>159</sup> As shown in Figure 5-7, a sharp band, corresponding to the fragment, was found on the gel image, indicating that the infectious pathogen part P had been successfully digested by the enzyme and lost its infection activity.

### **Performance of the Entire System with Pathogen Input**

After confirming the proper function of each step, we further tested the performance of entire system with the same pathogen input from *Bacillus anthracis* genome. First, 200 nM AM, 200 nM BM, 200 nM PG, 50 nM CP, 500 nM MB-R, 0.5 U/ $\mu$ L Phi29 and 250  $\mu$ M dNTP were mixed in the reaction buffer. To check the tolerance of AIS biomimicry, a small amount of P<sub>0</sub> (100 nM) was then added to the system, and fluorescence was monitored in real time. It can be observed from Figure 5-8A that the fluorescence kinetics is similar to that without P<sub>0</sub>, confirming that the system is not activated to produce antibody P\* when the amount of pathogen is under the level of immune tolerance (i.e., threshold). However, when another 250 nM P<sub>0</sub> were added to

the system to mimic the second exposure to the same pathogen, the fluorescence restoration showed much faster kinetics behavior, indicating that the production of a large amount of antibody P\* and successful trigger of immune response. In addition, the faster response in the second time exposure also confirmed the memory effect of this system for specific pathogen.

To further characterize the entire system, we used gel electrophoresis to analyze the final product. As shown in Figure 5-8B, only one sharp high molecular-weight band can be seen with the pathogen amount over the tolerance level (100nM +250 nM), which is consistent with the previous fluorescence result. To study the binding between pathogen and antibody in step 3, we modified P<sub>0</sub> with a fluorophore (FAM). After the generation of antibody P\*, the binding between P domain and P\* domain was confirmed to be present, as a high molecular weight band could be seen in the FITC imaging channel (Fig. 5-8C). With further addition of more FAM-labeled P<sub>0</sub> (from 1 time to 250 times excess) to the system, the hybridization between P and generated antibody P\* was still observed. For instance, as shown in lane 3 of Figure 5-8C, 2 μM more excess P<sub>0</sub> can be almost completely bound by the originally generated antibody P\*, indicating that the capacity of AIS mimicry is sufficient to quench excess pathogen due to the rolling cycle amplification reaction. Finally, we incubated SsPI restriction enzyme (2U/μL) together with all other components (200 nM AM, 200 nM BM, 200 nM PG, 50 nM CP, 0.5 U/ μL Phi29 and 250 μM dNTP). After introducing excess concentration of FAM-labeled P<sub>0</sub> (500 nM) that overcome the immune tolerance of the system, one major fragment band (56 nt) was discovered under the FITC channel, leading to the

conclusion that the pathogen was successfully digested by our AIS mimicry system (Fig. 5-8D, Fig. 5-9).

As discussed before, the specificity of the system mainly depends on the restriction enzyme and sequence of CP. To further confirm this AIS mimicry is specific for pathogen *Bacillus anthracis*, we challenge it by introducing a piece of FAM-labeled severe acute respiratory syndrome coronavirus genomic DNA (SARS) in pathogen input (P-SARS). Since there is no hybridization between P-SARS and generated antibody P\* only specific for pathogen *Bacillus anthracis*, no obvious high-molecular weight fluorescent band appeared with either small or large amount of P-SARS under FITC channel, indicating the specificity of this biomimicry system (Fig. 5-8E).

### **Conclusions**

Building on the rich knowledge of DNA and DNA-enzyme cascade reaction, we successfully erected an artificial DNA biomimicry network for AIS and also experimentally demonstrate its dynamic behaviors to emulate the real AIS. This is the first artificial molecular system based on DNA-DNA and DNA-protein interactions for mimicking the vertebrate host adaptive immune system. On a fundamental level, our approach provided a general design principle by using DNA for mimicking the naturally occurring system, including modularity, parameter sensitivity and kinetics from bottom-up. A more practical aspect of such artificial mimicry system, which reproduces some of the essential features of biological networks, may include the driving force for biological and biomedical applications regarding the potential of reaction networks, such as the guide for the search for an HIV vaccine.<sup>160</sup> Finally, the molecular programming strategy reported here could impact the development of powerful autonomous biocomputers or theranostic nanodevices.

Table 5-1. Sequences used in the AIS mimicry system

	Sequences (5'-3')
P <sub>0</sub>	GGATTATTGTTAAATATTGATAAGGATAACATTTCTCCAAC TAACTTACGTCCCTCATTCAATACCCTACG-InvT
P <sub>0</sub> -FAM	GGATTATTGTTAAATATTGATAAGGATAACATTTCTCCAAC TAACTTACGTCCCTCATTCAATACCCTACG-FAM-InvT
P-SARS	ATAATACTGCGTCTTGGTTCACAGCAACATTTCTCCAAC AACTTACGTCCCT-CATTCAATACCCTACG-InvT
AM <sub>1</sub>	AGACGTAGGGTATTGAATGAGGGATGTAA-InvT
TM	CATTCAATACCCTACGTCTCCA-InvT
BM <sub>1</sub>	TGGAGACGTAGGGTATTGAATGAGGGACGTAAGTTAGTT GGAGAAATGTT-InvT
BM <sub>2</sub> or PI	CCACTCTACCATAACATT TCTCCAACCTAACTTACGT-InvT
BM <sub>3</sub> or W1	CCCTCATTCAATACCCTACG-InvT
PG <sub>1</sub>	TGGAGAGTTGTTATGGTAGAGTGG
PG <sub>1</sub> -DAB	TGGAGAGTTGTTATGGTAGAGTGG-DAB
PG <sub>1</sub> -DAB-1mis	TGGAGAGATGTTATGGTAGAGTGG-DAB
PG <sub>1</sub> -DAB-2mis	TGGAGAAATGTTATGGTAGAGTGG-DAB
PG <sub>2</sub>	CCACTCTACCATAACACA
PG <sub>2</sub> -FAM	FAM-CCACTCTACCATAACACA
MB-R	FAM-CCGAGGGATTTTTTTTTTTGGATCTCGG-DAB
CP	PhosGTAGAGTGGGGATTATTGTTAAATATTGATAAGGATT TTTTTTTTTGGATTATTGTTAAATATTGATAAGGATTGTGTT ATG

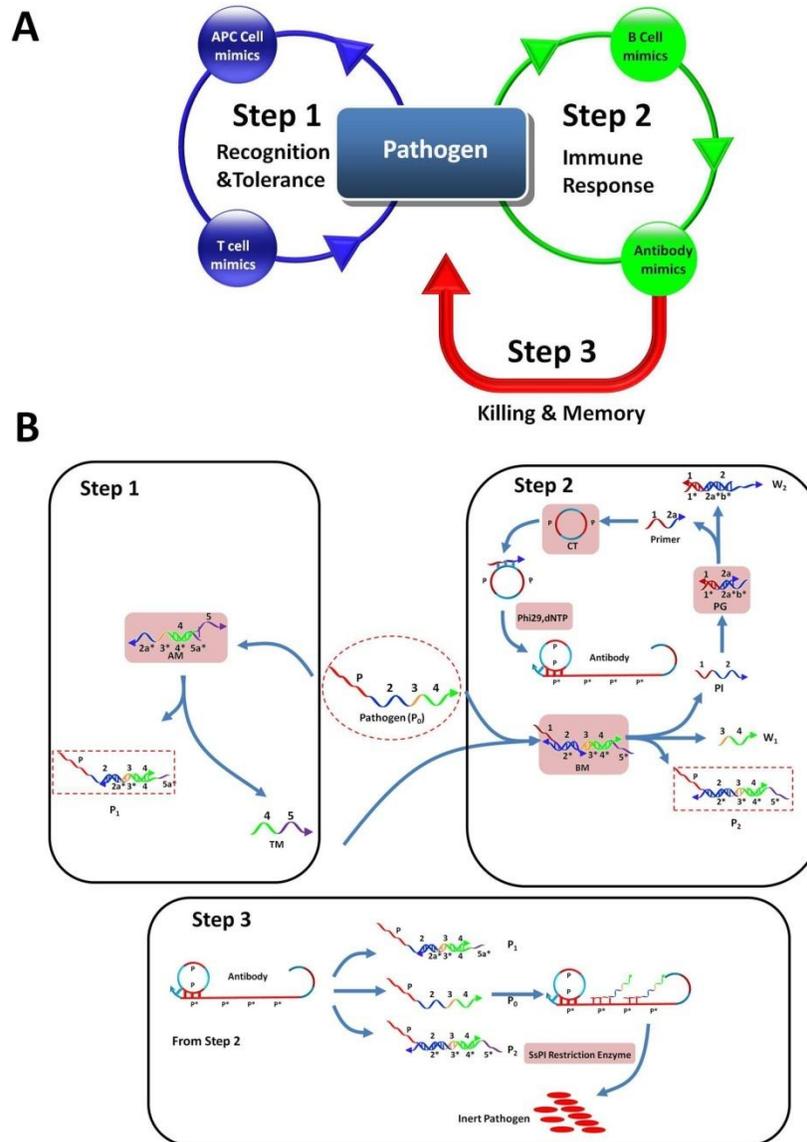


Figure 5-1. Working principle of AIS biomimicry system. A) Diagram illustrating system. The system consists of three steps: Recognition & Tolerance, Immune Response and Killing & Memory, which can be programmed by autonomous immunomimicking response with pathogen ( $P_0$ ) input. B) Working scheme for AIS mimicry system, driven by a series of DNA displacement and DNA-enzyme reactions. Colored lines indicate DNA strands with different domains. All  $x$  domains are complementary to  $x^*$ ;  $P_0$  is the pathogen sequence which possesses infection ability in ssDNA form. AM, BM, and PG are initially present as duplex components, along with circular DNA template CT and two functional enzymes of Phi29 polymerase and SsPI restriction enzyme.

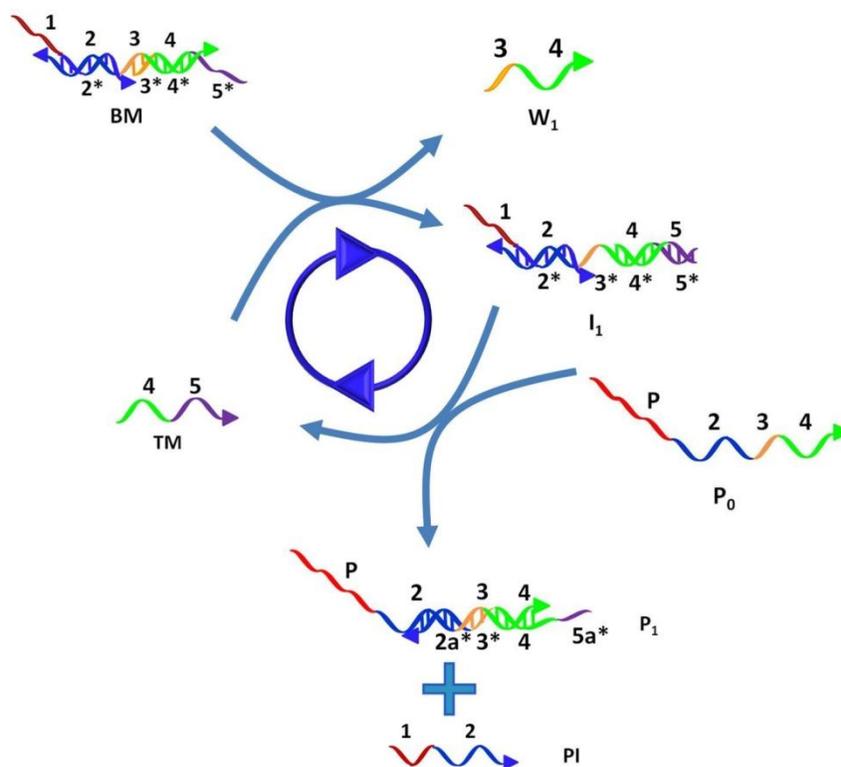


Figure 5-2. The proposed entropy-driven catalytic pathway.<sup>49</sup> TM first bind with BM through a 4-nt toehold region (domain 5), and form intermediate  $I_1$  and ssDNA  $W_1$  (34).  $P_0$  can bind with  $I_1$  by a newly formed 4-nt toehold (domain 3) and displace TM and ssDNA PI (12). The hybridization reaction is driven forward thermodynamically by the entropic gain of the liberated molecules.

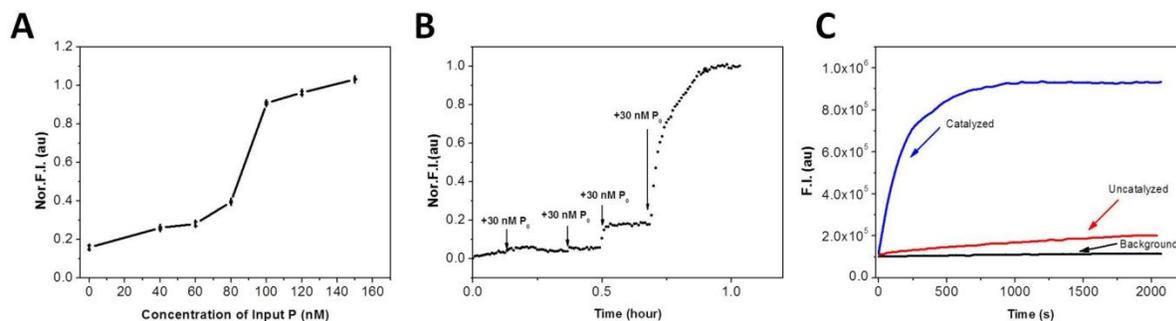


Figure 5-3. Fluorescence studies of the reaction priority between steps 1 and 2, as well as the catalytic effect of TM on the reaction of BM and P<sub>0</sub>. A) Plot of fluorescence restoration of the system with 80 nM AM, 100 nM BM and 100nM FAM/DAB-labeled PG with different concentrations of P<sub>0</sub>. B) Kinetics experiments of the system with 80 nM AM, 100 nM BM and 100 nM FAM/DAB-labeled PG. 30 nM P<sub>0</sub> were added to the buffer separately at each time point. C) Kinetics experiments to study the catalytic effect of TM. The black curve represents background kinetics with 80 nM AM, 100 nM BM, 100nM FAM/DAB-labeled PG and no P<sub>0</sub>. The red curve shows the uncatalyzed reaction kinetics with 100 nM BM, 100 nM FAM/DAB-labeled PG and 150 nM P<sub>0</sub>. The blue curve exhibits the catalyzed reaction kinetics with 80 nM AM, 100 nM BM, 100nM FAM/DAB-labeled PG and 150 nM P<sub>0</sub>. Experiments were performed at 25°C in 50 mM Tris-HCl buffer containing 10 mM MgCl<sub>2</sub>.

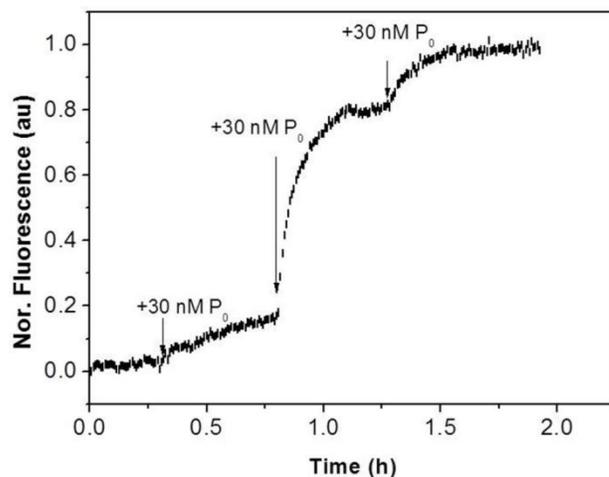


Figure 5-4. Fluorescence kinetics of the system with different P<sub>0</sub> additions. 30 nM P<sub>0</sub> was added to the buffer with 50 nM AM, 50 nM BM and 50 nM PG separately in each time.

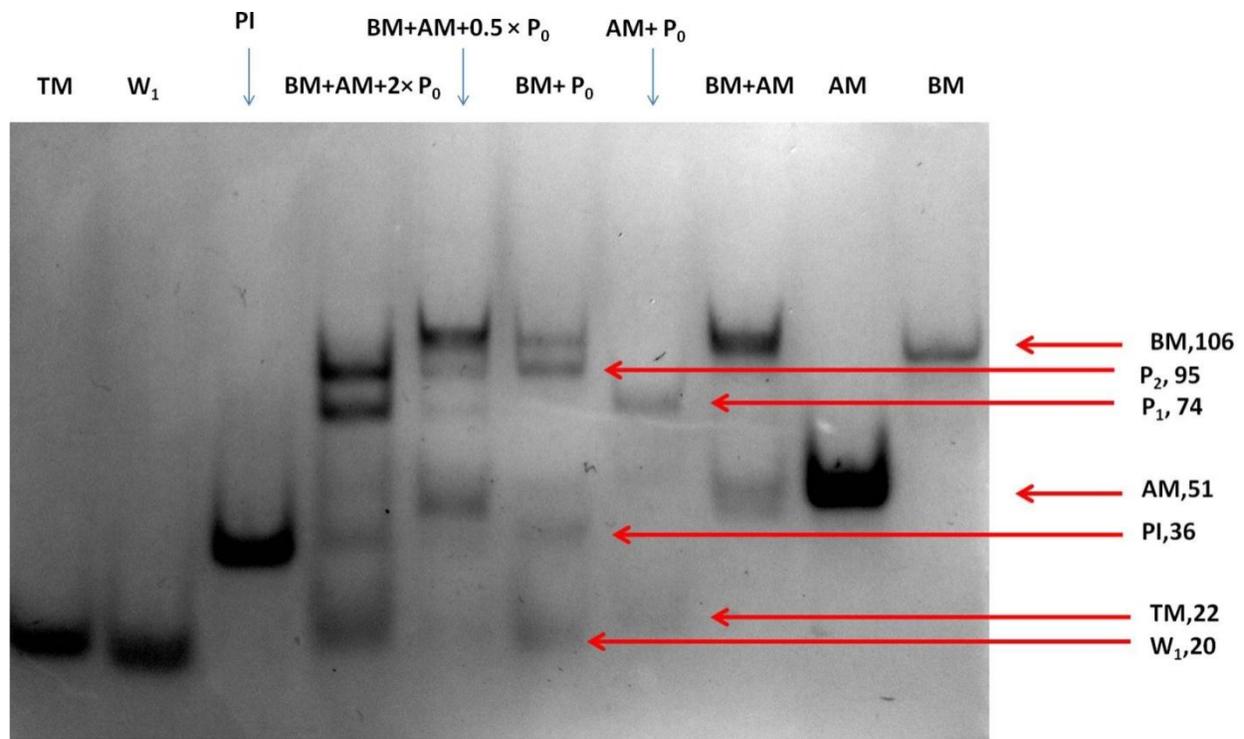


Figure 5-5. Analysis by PAGE (10% native gel) of the reaction pathway in step 1 and 2. Here, to fully separate DNA bands with similar molecular weight, P<sub>0</sub> sequence was modified by eliminating the *Bacillus anthracis* genomic sequence part (P: 28 bp). The products and their base numbers are indicated on right.

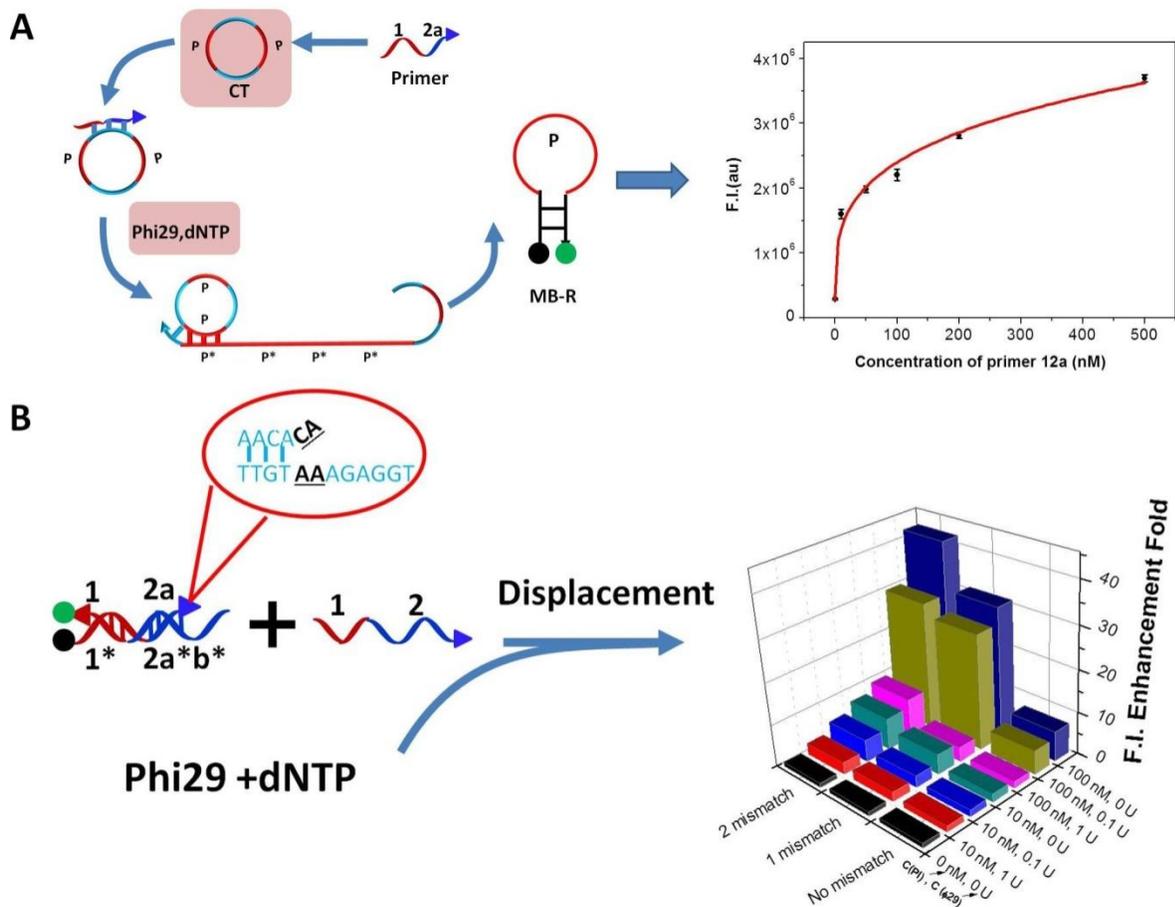


Figure 5-6. A) Scheme of primer-assisted RCA reaction and plot of fluorescence restoration versus different concentrations of primer 12a. B) Scheme of the design of mismatch points on PG and result of fluorescence enhancement fold versus different concentrations of ssDNA PI and Phi29. Experiments were performed at 25°C in 15mM Tris-HCl buffer containing 12.5 mM MgCl<sub>2</sub>, 10mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and 4mM DTT.

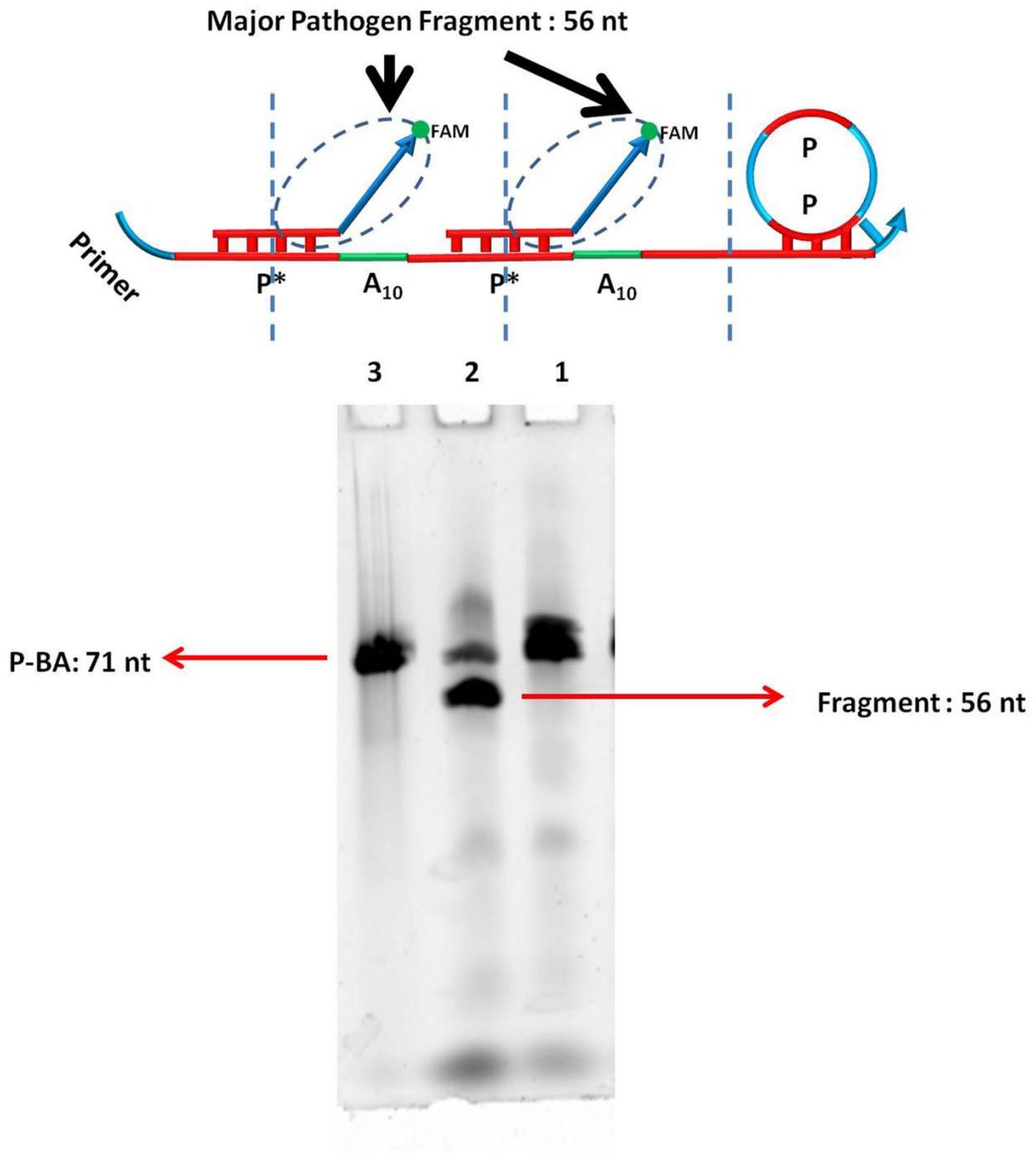


Figure 5-7. Scheme of the primer-initiated RCA product cut by SsPI restriction enzyme and their analysis by PAGE (8% denatured gel at 4 °C). Gel image was taken under FITC channel. Lane 1: RCA product + 1  $\mu$ M FAM-labeled P-SARS+2U/ $\mu$ L SsPI. Lane 2: RCA product + 1  $\mu$ M FAM-labeled P<sub>0</sub> +2U/ $\mu$ L SsPI. Lane3: FAM labeled P<sub>0</sub>.

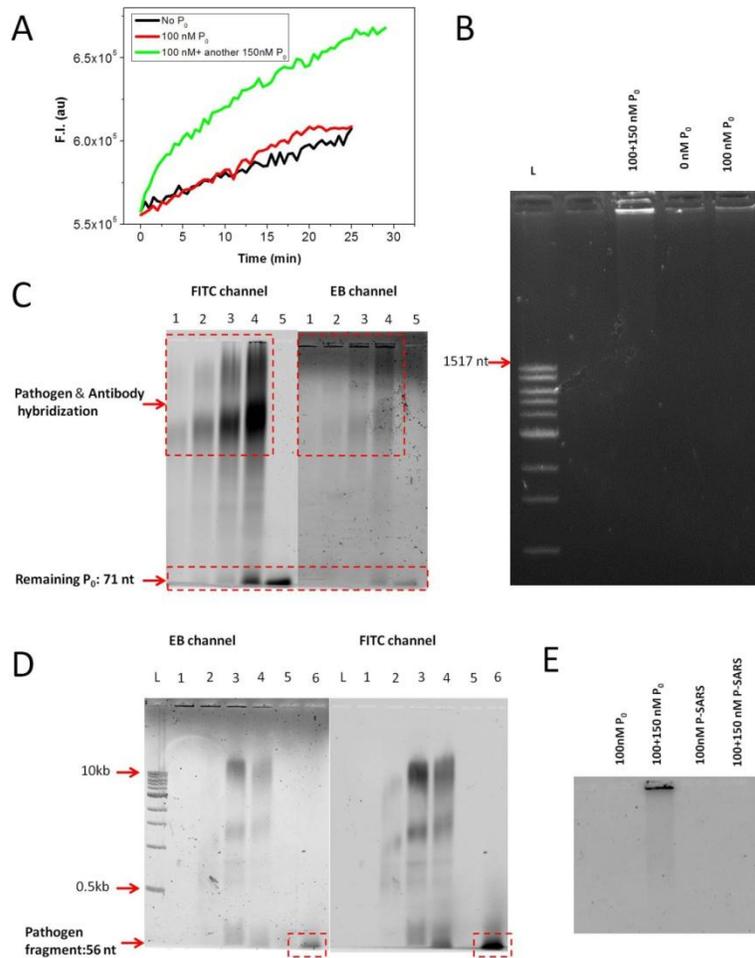


Figure 5-8. Experimental results of the entire system with pathogen input. A) Fluorescence kinetics of the antibody  $P^*$  generation with different concentrations of  $P_0$ . B) Analysis of the antibody  $P^*$  generation by agarose gel (1.5%). L: 100 bp ladder. The image was taken under EB channel. C) Analysis of the capacity of the system by agarose gel (1%). 200 nM AM, 200 nM BM, 200 nM PG, 50 nM CP, 0.5 U/  $\mu$ L Phi29 and 250  $\mu$ M dNTP were mixed in the reaction buffer followed by adding 350 nM FAM-labeled  $P_0$  and incubated for 1 hour. Afterwards, different amounts of  $P_0$  were added to mixture to test the binding capacity of the system for excess  $P_0$ . Lane 1: mixture +200 nM FAM-labeled  $P_0$ , lane 2: mixture +1  $\mu$ M FAM-labeled  $P_0$ , lane 3: mixture +2  $\mu$ M FAM-labeled  $P_0$ , lane 4: mixture +5  $\mu$ M FAM-labeled  $P_0$ , lane 5:  $P_0$  only. D) Analysis of the reactions in the entire system by agarose gel (1%). M = mixture of 200 nM AM, 200 nM BM, 200 nM PG, 50 nM CP, 0.5 U/  $\mu$ L Phi29 and 250  $\mu$ M dNTP. Incubation time = 1 hour. Lane 1: M+0 nM FAM-labeled  $P_0$ , lane 2: M+100nM FAM-labeled  $P_0$ , lane 3: M+500nM FAM-labeled  $P_0$ , lane 4: M+300nM FAM-labeled  $P_0$ , lane 5: M+100nM FAM-labeled  $P_0$ + 2U/ $\mu$ L SsPI, lane 6: M+500nM FAM-labeled  $P_0$ + 2U/ $\mu$ L SsPI, L: 1kb ladder. E) Agarose gel analysis of the system specificity (1.5%). The image was taken under FITC channel. FITC channel: EX=488nm, EM=512nm. EB channel: EX=488nm, EM=623nm.

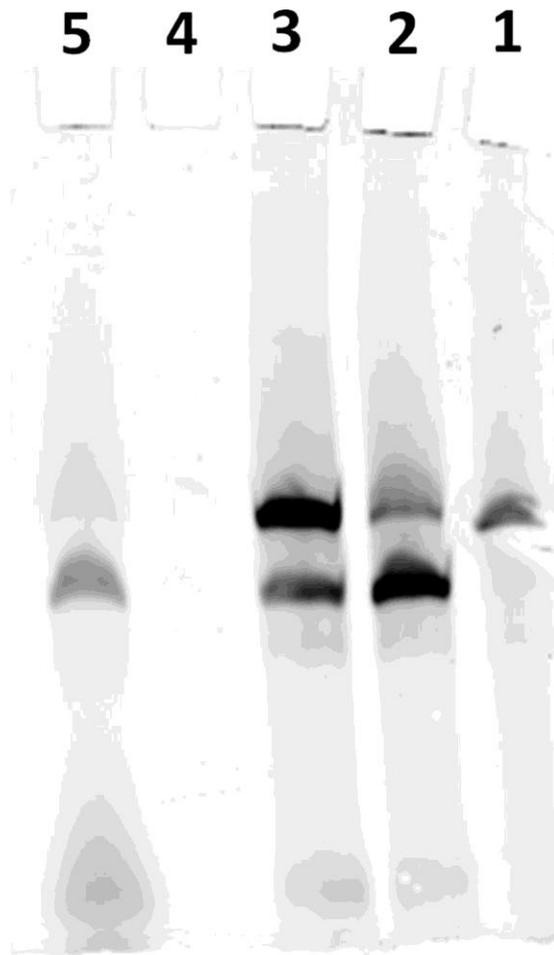


Figure 5-9. Gel analysis by PAGE (8% denatured gel) for the entire AIS mimicry system with SsPI restriction enzyme. Gel image was taken under FITC channel. M = mixture of 200 nM AM, 200 nM BM, 200 nM PG, 50 nM CP, 0.5 U/  $\mu$ L Phi29 and 250  $\mu$ M dNTP. Lane 1: FAM-labeled  $P_0$ . Lane 2: M + 1  $\mu$ M FAM-labeled  $P_0$ + 2U/ $\mu$ L SsPI for 2 hours. Lane3: M + 1  $\mu$ M FAM-labeled  $P_0$ + 2U/ $\mu$ L SsPI for 15 minutes. Lane 4: M only. Lane 5: 200 nM primer 12a+ 50 nM CP + 0.5 U/  $\mu$ L Phi29 and 250  $\mu$ M dNTP+1  $\mu$ M FAM-labeled  $P_0$ +2U/ $\mu$ L SsPI for 2 hours.

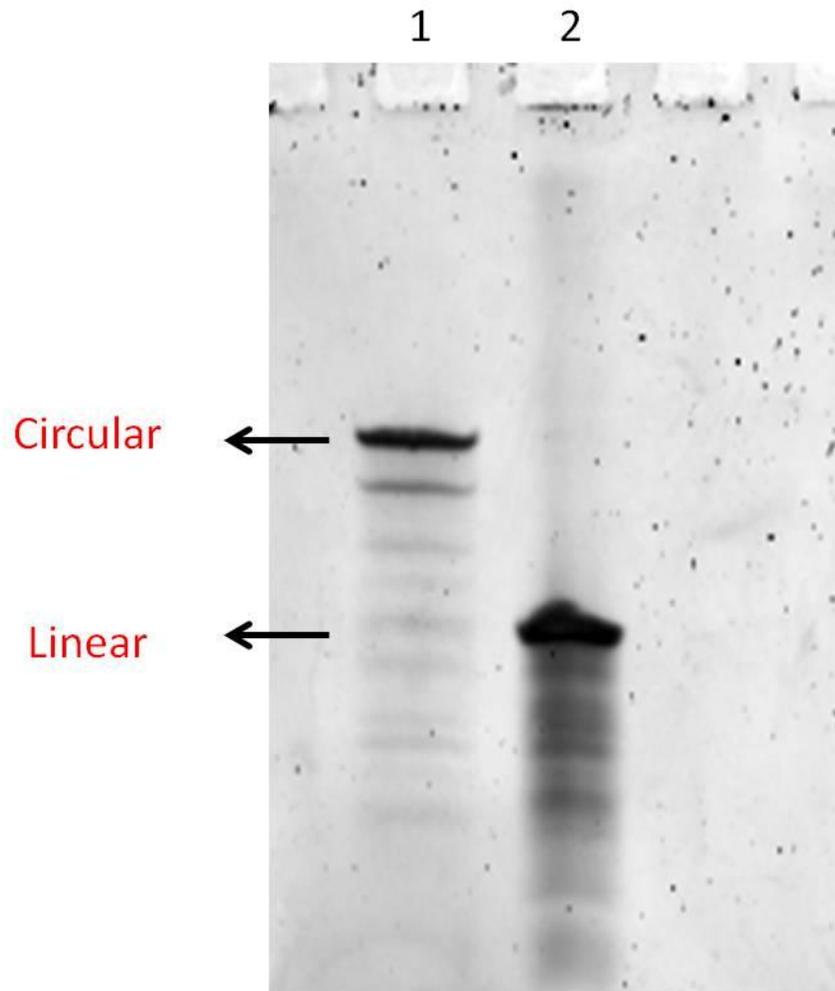


Figure 5-10. Analysis by PAGE (8% denatured gel) of the preparation of circular template (CP). Lane1: After ligation. Lane 2: Before ligation.

## CHAPTER 6 FUTURE DIRECTIONS AND CONCLUSIONS

### **Conclusions**

This thesis has reported the development of “smart” DNA systems for intelligent therapy on the molecular level via molecular engineering. Specifically, to qualify as “smart,” therapeutic systems should have the following properties: 1) specific recognition of disease-related biomarkers for sensitive diagnosis and signal amplification for effective therapy; 2) precise control of drug release with temporal and spatial resolution for effective therapy; 3) inhibition of drug side effects for rapid recovery; and 4) autonomous and programmable implementation with self-computation. To accomplish these objectives, DNA-based functional tools hold the greatest promise, due to their stability, programmability, and regulation of biological reactions. Therefore, this thesis has focused on engineering DNA molecules into smart molecular tools for intelligent therapeutic applications from the following four directions: 1) DNA aptamer-based system for accurate cancer cell identification and amplified cancer photo dynamic therapy; 2) photo-controllable DNA nanostructure platform for smart drug delivery; 3) DNA circuit for self-regulatory inhibition of drug side effects; and 4) DNA-based system for acquired immune system mimicry.

Chapter 2 demonstrated a DNA aptamer-based system for accurate identification of specific cancer cells and amplified photo dynamic therapy of cancers. The DNA circuit used in this strategy, composed of four functional modules ( $A_1$ ,  $A_2$ ,  $R_{12}$  and TDO5-C) totally made of DNA, can greatly amplify the singlet oxygen generation and selectively kill cancer cells. In particular, the DNA hairpin amplification circuit can be catalyzed by specifically designed nucleic acid sequences. Many nucleic acids,

including messenger RNA, microRNAs and small interfering RNAs, are important biomarkers for various diseases. If sequences for these biomolecules become available, the amplification hairpin DNA circuit can be designed to perform other biological and biomedical functions inside targeted disease cells with effective delivery methods. Second, increasing numbers of aptamers have been developed to target the membranes of a variety of cancer cell lines, thus establishing the universality of this DNA hairpin circuit for targeted and amplified therapy. Finally, as an application of DNA circuitry to biological cells, the prototype DNA circuit demonstrated here has the potential to enhance DNA technology with new insights and will broaden the utility of DNA circuits for applications in biology, biotechnology, and biomedicine.

The third chapter described a photoresponsive three-dimensional DNA nanostructure which can be used for smart drug delivery. DNA nanostructures capable of delivery of target molecules have attracted increasing interest because of their potential applications in biomedical and bioanalytical fields, such as drug delivery and molecular sensing. Compared to other nanomaterials, DNA is environmentally friendly, highly programmable and controllable. Therefore, it can be engineered into smart molecular delivery systems with precise temporal and spatial control on drug release, thus leading to the development of intelligent therapy. This research effort has been focused on the construction of a photoresponsive three-dimensional DNA drug delivery system. The working principle involves photoisomerization of azobenzene-incorporated DNA, wherein different wavelengths of light (UV and visible) can change the structure of the azobenzene compound and, hence, the shape of the DNA. This malleability is important to the design of this drug delivery system, because the trapping and release

of drug molecules can be controlled by light. Consequently, we have demonstrated that the utilization of light irradiation to control the shape of a DNA 3D nanostructure, thus enabling the release of cargos encapsulated inside this dynamic DNA nanostructure, as a smart drug delivery system with precise temporal and spatial resolution.

The fourth chapter reports a DNA-based logical circuit for self-regulatory inhibition of protein activity, which in turn can be used to autonomously control drug side effects. Effective inhibition of drug side effects is another significant issue in intelligent therapy. As the centers of homeostatic systems and the key regulators of an organism's behavior, proteins usually are the regulation targets of various drugs. Thrombin is a multifunctional protease that involved in a series of enzymatic and cellular reactions in the regulation of homeostasis. Impaired or overwhelming of thrombin formation can lead to various diseases. Normally impaired thrombin function is repaired by providing thrombin analogs or thrombin-induced drugs intravenously. However, precise dosage control of these drugs for individual patients has always been a difficult issue for physicians, as the unsuitable usage of these drugs may lead to severe side effects, such as blood clotting or even thrombosis. Therefore, to effectively inhibit the side effects of the thrombin-generating drugs and maintain the activity of thrombin at a constant level, a self-regulatory DNA logic circuit based on DNA-protein interactions was designed. This innovative design allows accurate threshold control, as well as autonomous, self-sustained and programmable manipulation of thrombin activity in vitro. This circuit can precisely sense the local enzymatic environment, specifically, the concentration of thrombin, and when it is excessively high, a coagulation inhibitor, i.e., a DNA anti-thrombin aptamer, is automatically released by a concentration-adjusted

circuit module. This prototype could lead to the development of novel DNA biochemical circuits to control the delivery of aptamer-based drugs in smart and personalized medicine, providing a more efficient and safer therapeutic strategy.

In the fifth chapter, for the first time, an artificial DNA biomimetic network was designed and constructed for vertebrate Adaptive Immune System based on DNA-DNA and DNA-protein interactions. When tested experimentally, the new network demonstrated its ability to mimic the dynamics of the vertebrate Adaptive Immune System. This strategy provides a general design principle by using DNA to mimic a naturally occurring system, including modularity, parameter sensitivity and kinetics from bottom-up. This model will also allow us to gain more insight into real reaction networks with concomitant translational applications, e.g., in the case of AIS, perhaps the search for an HIV vaccine.

### **Future Directions**

The successful completion of this dissertation research has demonstrated the potential of DNA molecular engineering in designing novel intelligent diagnostic and therapeutic systems. Building on this foundation, future research will expand the capability of those systems for further bioanalytical and biomedical applications. First, because of the complexity of cancer cell development, accurate identification of cancer subtypes is extremely significant for the successful treatment of cancers. Some sequence-specific DNA aptamer probes can effectively accomplish such target-specific recognition of cancer cells. However, single-aptamer recognition has not been sufficiently effective at identifying cancers in individual patients, due to the multiplicity of cancer biomarkers present on single cancer cells. In other words, some cancer subtypes express the same biomarker and cannot be identified by a single aptamer.

Therefore, engineering a DNA aptamer-based system that integrates multiple aptamer recognition and information processing ability could solve this problem, thus effectively decreasing the possibility of false diagnosis.

As a continuation of the work described in Chapter 2, a biaptameric system will be constructed to determine whether a cancer cell represents a specific cell line or a specific state. The cancer cells will exhibit fluorescence only when both of the two biomarkers are present, thereby improving the diagnosis accuracy. In addition, to identify the cancer cells more sensitively, an amplification strategy will be introduced via DNA molecular engineering. The ultimate goal of this project is applying this DNA aptamer system to clinical diagnosis and profiling individuals' cancer subtypes.

As a follow-up to the photo-responsive 3D DNA nanostructure described in Chapter 3, this nanocage will be used to trap drug molecules and construct an intelligent delivery drug platform. To test this idea, rhodamine molecules will be encapsulated as a model cargo inside the DNA nanocage. In accordance with the working principle, under UV light, the rhodamine will be trapped inside the DNA cages, but under visible irradiation, they will be released. Afterwards, an anticancer drug molecule, doxorubicin, will be used with this platform for intelligent cancer therapy.

The threshold-controlled DNA nanocircuit described in Chapter 4 achieved autonomous, self-sustained and programmable manipulation of the catalytic activity of thrombin. However, the circuit works only in the test tube in its current state. Future plans include molding the individual DNA strands onto a DNA scaffold to build a complete system and facilitate the coordinated operation *in vivo*. If that goal is achieved, this system can be used as an anticoagulation drug for thrombosis-related diseases.

## LIST OF REFERENCES

1. Tyagi, S. & Kramer, F.R. Molecular beacons: Probes that fluoresce upon hybridization. *Nature Biotechnology* **14**, 303-308 (1996).
2. Li, J.J., Fang, X., Schuster, S.M. & Tan, W. Molecular Beacons: A Novel Approach to Detect Protein - DNA Interactions. *Angew Chem Int Ed Engl* **39**, 1049-1052 (2000).
3. Shangguan, D., Cao, Z., Meng, L., Mallikaratchy, P. et al. Cell-specific aptamer probes for membrane protein elucidation in cancer cells. *J Proteome Res* **7**, 2133-2139 (2008).
4. Seeman, N.C. DNA in a material world. *Nature* **421**, 427-431 (2003).
5. Adleman, L.M. Molecular Computation of Solutions to Combinatorial Problems. *Science* **266**, 1021-1024 (1994).
6. Watson, J.D. & Crick, F.H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738 (1953).
7. Arnott, S. Principles of Nucleic-Acid Structure - Saenger, W. *Nature* **312**, 174-174 (1984).
8. Beaucage, S.L. & Iyer, R.P. Advances in the Synthesis of Oligonucleotides by the Phosphoramidite Approach. *Tetrahedron* **48**, 2223-2311 (1992).
9. Beaucage, S.L. & Iyer, R.P. The Synthesis of Modified Oligonucleotides by the Phosphoramidite Approach and Their Applications. *Tetrahedron* **49**, 6123-6194 (1993).
10. Ellington, A.D. & Szostak, J.W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822 (1990).
11. Hermann, T. & Patel, D.J. Adaptive recognition by nucleic acid aptamers. *Science* **287**, 820-825 (2000).
12. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510 (1990).
13. Shangguan, D., Li, Y., Tang, Z., Cao, Z.C. et al. Aptamers evolved from live cells as effective molecular probes for cancer study. *Proc Natl Acad Sci U S A* **103**, 11838-11843 (2006).
14. Wang, K., You, M., Chen, Y., Han, D. et al. Self-assembly of a bifunctional DNA carrier for drug delivery. *Angew Chem Int Ed Engl* **50**, 6098-6101.

15. Cho, E.J., Yang, L., Levy, M. & Ellington, A.D. Using a deoxyribozyme ligase and rolling circle amplification to detect a non-nucleic acid analyte, ATP. *J Am Chem Soc* **127**, 2022-2023 (2005).
16. Yang, C.J., Jockusch, S., Vicens, M., Turro, N.J. & Tan, W. Light-switching excimer probes for rapid protein monitoring in complex biological fluids. *Proc Natl Acad Sci U S A* **102**, 17278-17283 (2005).
17. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B. et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350-1354 (1985).
18. Shangguan, D., Li, Y., Tang, Z., Cao, Z. et al. Aptamers evolved from live cells as effective molecular probes for cancer study. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11838-11843 (2006).
19. Shangguan, D., Meng, L., Cao, Z.C., Xiao, Z. et al. Identification of liver cancer-specific aptamers using whole live cells. *Anal Chem* **80**, 721-728 (2008).
20. Sefah, K., Tang, Z., Shangguan, D., Chen, H. et al. Molecular recognition of acute myeloid leukemia using aptamers. *Leukemia* **23**, 235-244 (2009).
21. Chen, H.W., Medley, C.D., Sefah, K., Shangguan, D. et al. Molecular recognition of small-cell lung cancer cells using aptamers. *Chemmedchem* **3**, 991-1001 (2008).
22. Zhao, Z., Xu, L., Shi, X., Tan, W. et al. Recognition of subtype non-small cell lung cancer by DNA aptamers selected from living cells. *Analyst* **134**, 1808-1814 (2009).
23. Simaey, D.V., López-Colón, D., Sefah, K., Sutphen, R. et al. Study of the Molecular Recognition of Aptamers Selected through Ovarian Cancer Cell-SELEX. *Plos One* **5** (2010).
24. Mallikaratchy, P., Tang, Z., Kwame, S., Meng, L. et al. Aptamer Directly Evolved from Live Cells Recognizes Membrane Bound Immunoglobulin Heavy Mu Chain in Burkitt's Lymphoma Cells. *Molecular & Cellular Proteomics* **6**, 2230-2238 (2007).
25. Sefah, K., Meng, L., Lopez-Colon, D., Jimenez, E. et al. DNA Aptamers as Molecular Probes for Colorectal Cancer Study. *Plos One* **5** (2010).
26. Zhang, K., Sefah, K., Tang, L., Zhao, Z. et al. A Novel Aptamer Developed for Breast Cancer Cell Internalization. *Chemmedchem* **7**, 79-84 (2012).
27. Liu, Y., Lin, C.X., Li, H.Y. & Yan, H. Protein nanoarrays - Aptamer-directed self-assembly of protein arrays on a DNA nanostructure. *Angew Chem Int Edit* **44**, 4333-4338 (2005).

28. LaBean, T.H., Yan, H., Kopatsch, J., Liu, F.R. et al. Construction, analysis, ligation, and self-assembly of DNA triple crossover complexes. *Journal of the American Chemical Society* **122**, 1848-1860 (2000).
29. Winfree, E., Liu, F., Wenzler, L.A. & Seeman, N.C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394**, 539-544 (1998).
30. Liu, D., Park, S.H., Reif, J.H. & LaBean, T.H. DNA nanotubes self-assembled from triple-crossover tiles as templates for conductive nanowires. *P Natl Acad Sci USA* **101**, 717-722 (2004).
31. Rothemund, P.W.K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302 (2006).
32. Alivisatos, A.P., Johnsson, K.P., Peng, X.G., Wilson, T.E. et al. Organization of 'nanocrystal molecules' using DNA. *Nature* **382**, 609-611 (1996).
33. He, Y., Tian, Y., Ribbe, A.E. & Mao, C.D. Antibody nanoarrays with a pitch of similar to 20 nanometers. *Journal of the American Chemical Society* **128**, 12664-12665 (2006).
34. Sharma, J., Chhabra, R., Liu, Y., Ke, Y.G. & Yan, H. DNA-templated self-assembly of two-dimensional and periodical gold nanoparticle arrays. *Angew Chem Int Edit* **45**, 730-735 (2006).
35. Shih, W.M., Quispe, J.D. & Joyce, G.F. A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature* **427**, 618-621 (2004).
36. He, Y., Ye, T., Su, M., Zhang, C. et al. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452**, 198-U141 (2008).
37. Goodman, R.P., Schaap, I.A.T., Tardin, C.F., Erben, C.M. et al. Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. *Science* **310**, 1661-1665 (2005).
38. Andersen, E.S., Dong, M., Nielsen, M.M., Jahn, K. et al. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459**, 73-U75 (2009).
39. Douglas, S.M., Bachelet, I. & Church, G.M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **335**, 831-834 (2012).
40. Liu, Z.F., Hashimoto, K. & Fujishima, A. Photoelectrochemical Information-Storage Using an Azobenzene Derivative. *Nature* **347**, 658-660 (1990).
41. Asanuma, H., Ito, T., Yoshida, T., Liang, X. & Komiyama, M. Photoregulation of the Formation and Dissociation of a DNA Duplex by Using the cis-trans Isomerization of Azobenzene. *Angew Chem Int Ed Engl* **38**, 2393-2395 (1999).

42. Liang, X., Asanuma, H., Kashida, H., Takasu, A. et al. NMR study on the photoresponsive DNA tethering an azobenzene. Assignment of the absolute configuration of two diastereomers and structure determination of their duplexes in the trans-form. *J Am Chem Soc* **125**, 16408-16415 (2003).
43. Liang, X., Mochizuki, T. & Asanuma, H. A supra-photoswitch involving sandwiched DNA base pairs and azobenzenes for light-driven nanostructures and nanodevices. *Small* **5**, 1761-1768 (2009).
44. Kashida, H., Takatsu, T., Fujii, T., Sekiguchi, K. et al. In-stem molecular beacon containing a pseudo base pair of threoninol nucleotides for the removal of background emission. *Angew Chem Int Ed Engl* **48**, 7044-7047 (2009).
45. Ball, P. Chemistry meets computing. *Nature* **406**, 118-120 (2000).
46. Seelig, G., Soloveichik, D., Zhang, D.Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585-1588 (2006).
47. Yurke, B., Turberfield, A.J., Mills, A.P., Simmel, F.C. & Neumann, J.L. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605-608 (2000).
48. Zhang, D.Y. & Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *Journal of the American Chemical Society* **131**, 17303-17314 (2009).
49. Zhang, D.Y., Turberfield, A.J., Yurke, B. & Winfree, E. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318**, 1121-1125 (2007).
50. Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368-372 (2011).
51. Qian, L. & Winfree, E. Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **332**, 1196-1201 (2011).
52. Dirks, R.M. & Pierce, N.A. Triggered amplification by hybridization chain reaction. *P Natl Acad Sci USA* **101**, 15275-15278 (2004).
53. Zhang, D.Y., Turberfield, A.J., Yurke, B. & Winfree, E. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318**, 1121-1125 (2007).
54. Han, D., Zhu, Z., Wu, C.C., Peng, L. et al. A Logical Molecular Circuit for Programmable and Autonomous Regulation of Protein Activity Using DNA Aptamer-Protein Interactions. *Journal of the American Chemical Society* **134**, 20797-20804 (2012).

55. Venkataraman, S., Dirks, R.M., Ueda, C.T. & Pierce, N.A. Selective cell death mediated by small conditional RNAs (Retracted article. See vol. 110, pg. 384, 2013). *P Natl Acad Sci USA* **107**, 16777-16782 (2010).
56. Friedland, A.E., Lu, T.K., Wang, X., Shi, D. et al. Synthetic Gene Networks That Count. *Science* **324**, 1199-1202 (2009).
57. Kramer, B.P., Viretta, A.U., El Baba, M.D., Auel, D. et al. An engineered epigenetic transgene switch in mammalian cells. *Nature Biotechnology* **22**, 867-870 (2004).
58. Kim, J., White, K.S. & Winfree, E. Construction of an in vitro bistable circuit from synthetic transcriptional switches. *Molecular Systems Biology* **2** (2006).
59. Fujii, T. & Rondelez, Y. Predator-Prey Molecular Ecosystems. *Acs Nano* **7**, 27-34 (2013).
60. Padirac, A., Fujii, T., Estevez-Torres, A. & Rondelez, Y. Spatial Waves in Synthetic Biochemical Networks. *J Am Chem Soc* (2013).
61. Fu, T.J. & Seeman, N.C. DNA Double-Crossover Molecules. *Biochemistry* **32**, 3211-3220 (1993).
62. Dougherty, T.J., Gomer, C.J., Henderson, B.W., Jori, G. et al. Photodynamic therapy. *J Natl Cancer Inst* **90**, 889-905 (1998).
63. Bugaj, A.M. Targeted photodynamic therapy - a promising strategy of tumor treatment. *Photoch Photobio Sci* **10**, 1097-1109 (2011).
64. Mallikaratchy, P., Tang, Z. & Tan, W. Cell specific aptamer-photosensitizer conjugates as a molecular tool in photodynamic therapy. *ChemMedChem* **3**, 425-428 (2008).
65. Yang, X., Huang, J., Wang, K., Li, W. et al. Angiogenin-mediated photosensitizer-aptamer conjugate for photodynamic therapy. *ChemMedChem* **6**, 1778-1780 (2011).
66. Wang, J., Zhu, G., You, M., Song, E. et al. Assembly of aptamer switch probes and photosensitizer on gold nanorods for targeted photothermal and photodynamic cancer therapy. *Acs Nano* **6**, 5070-5077 (2012).
67. Zheng, G., Chen, J., Stefflova, K., Jarvi, M. et al. Photodynamic molecular beacon as an activatable photosensitizer based on protease-controlled singlet oxygen quenching and activation. *P Natl Acad Sci USA* **104**, 8989-8994 (2007).
68. McDonnell, S.O., Hall, M.J., Allen, L.T., Byrne, A. et al. Supramolecular photonic therapeutic agents. *Journal of the American Chemical Society* **127**, 16360-16361 (2005).

69. Tang, Z.W., Zhu, Z., Mallikaratchy, P., Yang, R.H. et al. Aptamer-Target Binding Triggered Molecular Mediation of Singlet Oxygen Generation. *Chem-Asian J* **5**, 783-786 (2010).
70. Zhu, Z., Tang, Z.W., Phillips, J.A., Yang, R.H. et al. Regulation of singlet oxygen generation using single-walled carbon nanotubes. *Journal of the American Chemical Society* **130**, 10856-+ (2008).
71. Clo, E., Snyder, J.W., Voigt, N.V., Ogilby, P.R. & Gothelf, K.V. DNA-programmed control of photosensitized singlet oxygen production. *Journal of the American Chemical Society* **128**, 4200-4201 (2006).
72. Chen, J., Stefflova, K., Niedre, M.J., Wilson, B.C. et al. Protease-triggered photosensitizing beacon based on singlet oxygen quenching and activation. *Journal of the American Chemical Society* **126**, 11450-11451 (2004).
73. Huang, J., Wu, Y.R., Chen, Y., Zhu, Z. et al. Pyrene-Excimer Probes Based on the Hybridization Chain Reaction for the Detection of Nucleic Acids in Complex Biological Fluids. *Angew Chem Int Edit* **50**, 401-404 (2011).
74. Yin, P., Choi, H.M.T., Calvert, C.R. & Pierce, N.A. Programming biomolecular self-assembly pathways. *Nature* **451**, 318-U314 (2008).
75. Li, B.L., Ellington, A.D. & Chen, X. Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Research* **39** (2011).
76. Tuerk, C. & Gold, L. Systematic Evolution of Ligands by Exponential Enrichment - Rna Ligands to Bacteriophage-T4 DNA-Polymerase. *Science* **249**, 505-510 (1990).
77. Robertson, D.L. & Joyce, G.F. Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* **344**, 467-468 (1990).
78. Huizenga, D.E. & Szostak, J.W. A DNA Aptamer That Binds Adenosine and Atp. *Biochemistry* **34**, 656-665 (1995).
79. Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H. & Toole, J.J. Selection of Single-Stranded-DNA Molecules That Bind and Inhibit Human Thrombin. *Nature* **355**, 564-566 (1992).
80. Shangguan, D., Li, Y., Tang, Z.W., Cao, Z.H.C. et al. Aptamers evolved from live cells as effective molecular probes for cancer study. *P Natl Acad Sci USA* **103**, 11838-11843 (2006).
81. Tang, Z.W., Shangguan, D., Wang, K.M., Shi, H. et al. Selection of aptamers for molecular recognition and characterization of cancer cells. *Anal Chem* **79**, 4900-4907 (2007).

82. Sefah, K., Tang, Z.W., Shangguan, D.H., Chen, H. et al. Molecular recognition of acute myeloid leukemia using aptamers. *Leukemia* **23**, 235-244 (2009).
83. Shangguan, D.H., Meng, L., Cao, Z.H.C., Xiao, Z.Y. et al. Identification of liver cancer-specific aptamers using whole live cells. *Anal Chem* **80**, 721-728 (2008).
84. Wang, L., Yang, C.J., Medley, C.D., Benner, S.A. & Tan, W. Locked nucleic acid molecular beacons. *J Am Chem Soc* **127**, 15664-15665 (2005).
85. Tang, Z., Mallikaratchy, P., Yang, R., Kim, Y. et al. Aptamer switch probe based on intramolecular displacement. *J Am Chem Soc* **130**, 11268-11269 (2008).
86. Kim, V.N. Small RNAs: classification, biogenesis, and function. *Mol Cells* **19**, 1-15 (2005).
87. Carthew, R.W. & Sontheimer, E.J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642-655 (2009).
88. Venkataraman, S., Dirks, R.M., Ueda, C.T. & Pierce, N.A. Selective cell death mediated by small conditional RNAs. *Proc Natl Acad Sci U S A* **107**, 16777-16782 (2010).
89. Seeman, N.C. From genes to machines: DNA nanomechanical devices. *Trends Biochem Sci* **30**, 119-125 (2005).
90. Seeman, N.C. Nucleic-Acid Junctions and Lattices. *Journal of Theoretical Biology* **99**, 237-247 (1982).
91. Malo, J., Mitchell, J.C., Venien-Bryan, C., Harris, J.R. et al. Engineering a 2D protein-DNA crystal. *Angewandte Chemie-International Edition* **44**, 3057-3061 (2005).
92. Tanaka, F., Mochizuki, T., Liang, X.G., Asanuma, H. et al. Robust and Photocontrollable DNA Capsules Using Azobenzenes. *Nano Letters* **10**, 3560-3565 (2010).
93. Seelig, G., Soloveichik, D., Zhang, D.Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585-1588 (2006).
94. Goodman, R.P., Heilemann, M., Doose, S., Erben, C.M. et al. Reconfigurable, braced, three-dimensional DNA nanostructures. *Nature Nanotechnology* **3**, 93-96 (2008).
95. Bath, J., Green, S.J. & Turberfield, A.J. A free-running DNA motor powered by a nicking enzyme. *Angewandte Chemie-International Edition* **44**, 4358-4361 (2005).

96. Liu, D.S. & Balasubramanian, S. A proton-fuelled DNA nanomachine. *Angewandte Chemie-International Edition* **42**, 5734-5736 (2003).
97. Simmel, F.C. & Dittmer, W.U. DNA nanodevices. *Small* **1**, 284-299 (2005).
98. Kumar, G.S. & Neckers, D.C. Photochemistry of Azobenzene-Containing Polymers. *Chemical Reviews* **89**, 1915-1925 (1989).
99. Asanuma, H., Ito, T., Yoshida, T., Liang, X.G. & Komiyama, M. Photoregulation of the formation and dissociation of a DNA duplex by using the cis-trans isomerization of azobenzene. *Angewandte Chemie-International Edition* **38**, 2393-2395 (1999).
100. Liang, X.G., Mochizuki, T. & Asanuma, H. A Supra-photoswitch Involving Sandwiched DNA Base Pairs and Azobenzenes for Light-Driven Nanostructures and Nanodevices. *Small* **5**, 1761-1768 (2009).
101. Asanuma, H., Liang, X., Nishioka, H., Matsunaga, D. et al. Synthesis of azobenzene-tethered DNA for reversible photo-regulation of DNA functions: hybridization and transcription. *Nature Protocols* **2**, 203-212 (2007).
102. Liang, X.G., Nishioka, H., Takenaka, N. & Asanuma, H. A DNA nanomachine powered by light irradiation. *ChemBiochem* **9**, 702-705 (2008).
103. Kang, H.Z., Liu, H.P., Phillips, J.A., Cao, Z.H. et al. Single-DNA Molecule Nanomotor Regulated by Photons. *Nano Lett* **9**, 2690-2696 (2009).
104. Lee, J.H., Wernette, D.P., Yigit, M.V., Liu, J. et al. Site-specific control of distances between gold nanoparticles using phosphorothioate anchors on DNA and a short bifunctional molecular fastener. *Angew Chem Int Edit* **46**, 9006-9010 (2007).
105. Luduena, R.F., Roach, M.C., Trcka, P.P. & Weintraub, S. N,N-Bis(Alpha-Iodoacetyl)-2,2'-Dithiobis(Ethylamine), a Reversible Crosslinking Reagent for Protein Sulfhydryl-Groups. *Anal Biochem* **117**, 76-80 (1981).
106. Deng, Z.X., Tian, Y., Lee, S.H., Ribbe, A.E. & Mao, C.D. DNA-encoded self-assembly of gold nanoparticles into one-dimensional arrays. *Angew Chem Int Edit* **44**, 3582-3585 (2005).
107. Goodman, R.P., Berry, R.M. & Turberfield, A.J. The single-step synthesis of a DNA tetrahedron. *Chem Commun (Camb)*, 1372-1373 (2004).
108. Zhang, D.Y. & Winfree, E. Robustness and modularity properties of a non-covalent DNA catalytic reaction. *Nucleic Acids Research* **38**, 4182-4197 (2010).
109. Ran, T., Kaplan, S. & Shapiro, E. Molecular implementation of simple logic programs. *Nat Nanotechnol* **4**, 642-648 (2009).

110. Beisel, C.L., Bayer, T.S., Hoff, K.G. & Smolke, C.D. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Molecular Systems Biology* **4** (2008).
111. Kolpashchikov, D.M. & Stojanovic, M.N. Boolean control of aptamer binding states. *Journal of the American Chemical Society* **127**, 11348-11351 (2005).
112. Stojanovic, M.N. & Kolpashchikov, D.M. Modular aptameric sensors. *Journal of the American Chemical Society* **126**, 9266-9270 (2004).
113. Benenson, Y., Gil, B., Ben-Dor, U., Adar, R. & Shapiro, E. An autonomous molecular computer for logical control of gene expression. *Nature* **429**, 423-429 (2004).
114. Gil, B., Kahan-Hanum, M., Skirtenko, N., Adar, R. & Shapiro, E. Detection of multiple disease indicators by an autonomous biomolecular computer. *Nano Lett* **11**, 2989-2996 (2011).
115. Saghatelian, A., Volcker, N.H., Guckian, K.M., Lin, V.S. & Ghadiri, M.R. DNA-based photonic logic gates: AND, NAND, and INHIBIT. *J Am Chem Soc* **125**, 346-347 (2003).
116. Win, M.N. & Smolke, C.D. Higher-order cellular information processing with synthetic RNA devices. *Science* **322**, 456-460 (2008).
117. Carothers, J.M., Goler, J.A., Juminaga, D. & Keasling, J.D. Model-driven engineering of RNA devices to quantitatively program gene expression. *Science* **334**, 1716-1719 (2011).
118. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science* **333**, 1307-1311 (2011).
119. Stricker, J., Cookson, S., Bennett, M.R., Mather, W.H. et al. A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516-519 (2008).
120. Freeman, M. Feedback control of intercellular signalling in development. *Nature* **408**, 313-319 (2000).
121. Culler, S.J., Hoff, K.G. & Smolke, C.D. Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. *Science* **330**, 1251-1255 (2010).
122. Vuyisich, M. & Beal, P.A. Controlling protein activity with ligand-regulated RNA aptamers. *Chem Biol* **9**, 907-913 (2002).

123. Lee, J.H., Canny, M.D., De Erkenez, A., Krilleke, D. et al. A therapeutic aptamer inhibits angiogenesis by specifically targeting the heparin binding domain of VEGF165. *Proc Natl Acad Sci U S A* **102**, 18902-18907 (2005).
124. Green, L.S., Jellinek, D., Jenison, R., Ostman, A. et al. Inhibitory DNA ligands to platelet-derived growth factor B-chain. *Biochemistry* **35**, 14413-14424 (1996).
125. Lee, J.F., Stovall, G.M. & Ellington, A.D. Aptamer therapeutics advance. *Curr Opin Chem Biol* **10**, 282-289 (2006).
126. Ni, X., Castanares, M., Mukherjee, A. & Lupold, S.E. Nucleic acid aptamers: clinical applications and promising new horizons. *Curr Med Chem* **18**, 4206-4214 (2011).
127. Hirsh, J. Choice of low molecular weight heparins. *Journal of Thrombosis and Haemostasis* **1**, 598-599 (2003).
128. Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H. & Toole, J.J. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **355**, 564-566 (1992).
129. Tasset, D.M., Kubik, M.F. & Steiner, W. Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J Mol Biol* **272**, 688-698 (1997).
130. Zhang, D.Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *J Am Chem Soc* **131**, 17303-17314 (2009).
131. Zhang, D.Y. & Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nat Chem* **3**, 103-113 (2011).
132. Nutiu, R. & Li, Y. Structure-switching signaling aptamers. *J Am Chem Soc* **125**, 4771-4778 (2003).
133. Wolberg, A.S. Thrombin generation and fibrin clot structure. *Blood Rev* **21**, 131-142 (2007).
134. Muller, J., Becher, T., Braunstein, J., Berdel, P. et al. Profiling of active thrombin in human blood by supramolecular complexes. *Angew Chem Int Ed Engl* **50**, 6075-6078 (2011).
135. Ramjee, M.K. The use of fluorogenic substrates to monitor thrombin generation for the analysis of plasma and whole blood coagulation. *Anal Biochem* **277**, 11-18 (2000).
136. Shiang, Y.C., Hsu, C.L., Huang, C.C. & Chang, H.T. Gold nanoparticles presenting hybridized self-assembled aptamers that exhibit enhanced inhibition of thrombin. *Angew Chem Int Ed Engl* **50**, 7660-7665 (2011).

137. Hsu, C.L., Chang, H.T., Chen, C.T., Wei, S.C. et al. Highly efficient control of thrombin activity by multivalent nanoparticles. *Chemistry* **17**, 10994-11000 (2011).
138. Franco, E., Friedrichs, E., Kim, J., Jungmann, R. et al. Timing molecular motion and production with a synthetic transcriptional clock. *Proc Natl Acad Sci U S A* **108**, E784-793 (2011).
139. Li, X., Shen, L., Zhang, D., Qi, H. et al. Electrochemical impedance spectroscopy for study of aptamer-thrombin interfacial interactions. *Biosens Bioelectron* **23**, 1624-1630 (2008).
140. Hood, L., Heath, J.R., Phelps, M.E. & Lin, B. Systems biology and new technologies enable predictive and preventative medicine. *Science* **306**, 640-643 (2004).
141. Choi, H.M., Chang, J.Y., Trinh le, A., Padilla, J.E. et al. Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat Biotechnol* **28**, 1208-1212 (2010).
142. Rothemund, P.W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302 (2006).
143. Douglas, S.M., Dietz, H., Liedl, T., Hogberg, B. et al. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414-418 (2009).
144. Bottomly, K. 1984: all idiotypes are equal, but some are more equal than others. *Immunol Rev* **79**, 45-61 (1984).
145. Krishnan, Y. & Simmel, F.C. Nucleic Acid Based Molecular Devices. *Angew Chem Int Edit* **50**, 3124-3156 (2011).
146. Simmel, F.C. Programming the Dynamics of Biochemical Reaction Networks. *Acs Nano* **7**, 6-10 (2013).
147. Johne, R., Muller, H., Rector, A., van Ranst, M. & Stevens, H. Rolling-circle amplification of viral DNA genomes using phi29 polymerase. *Trends Microbiol* **17**, 205-211 (2009).
148. Franco, E., Friedrichs, E., Kim, J., Jungmann, R. et al. Timing molecular motion and production with a synthetic transcriptional clock. *P Natl Acad Sci USA* **108**, E784-E793 (2011).
149. Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci U S A* **102**, 3581-3586 (2005).

150. Lou, C., Liu, X., Ni, M., Huang, Y. et al. Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch. *Mol Syst Biol* **6**, 350 (2010).
151. Padirac, A., Fujii, T. & Rondelez, Y. Bottom-up construction of in vitro switchable memories. *P Natl Acad Sci USA* **109**, E3212-E3220 (2012).
152. Zhang, D.Y. & Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nat Chem* **3**, 103-113 (2011).
153. Zhao, W.A., Ali, M.M., Brook, M.A. & Li, Y.F. Rolling circle amplification: Applications in nanotechnology and biodetection with functional nucleic acids. *Angew Chem Int Edit* **47**, 6330-6337 (2008).
154. Schweitzer, B., Wiltshire, S., Lambert, J., O'Malley, S. et al. Immunoassays with rolling circle DNA amplification: A versatile platform for ultrasensitive antigen detection. *P Natl Acad Sci USA* **97**, 10113-10119 (2000).
155. Fire, A. & Xu, S.Q. Rolling Replication of Short DNA Circles. *P Natl Acad Sci USA* **92**, 4641-4645 (1995).
156. Chen, X. & Ellington, A.D. Shaping up nucleic acid computation. *Curr Opin Biotech* **21**, 392-400 (2010).
157. Ortigao, J.F., Rosch, H., Selter, H., Frohlich, A. et al. Antisense effect of oligodeoxynucleotides with inverted terminal internucleotidic linkages: a minimal modification protecting against nucleolytic degradation. *Antisense Res Dev* **2**, 129-146 (1992).
158. Vandesande, J.H., Ramsing, N.B., Germann, M.W., Elhorst, W. et al. Parallel Stranded DNA. *Science* **241**, 551-557 (1988).
159. McManus, S.A. & Li, Y.F. Turning a Kinase Deoxyribozyme into a Sensor. *Journal of the American Chemical Society* **135**, 7181-7186 (2013).
160. Amara, R.R., Villinger, F., Altman, J.D., Lydy, S.L. et al. Control of a mucosal challenge and prevention of AIDS by a multiprotein DNA/MVA vaccine. *Science* **292**, 69-74 (2001).

## BIOGRAPHICAL SKETCH

Da Han was born in Heilongjiang, China. He spent his first 10 years in Qiqihaer, Heilongjiang. After elementary school education in 1998, he and his parents moved to a beautiful beach city, Qinhuangdao in Hebei province. In 2005, he attended Xiamen University, where he obtained his B.S. degree in chemistry in 2009. With the successful completion of all the necessary examinations for U.S. college graduate school application, he was accepted by Department of Chemistry in University of Florida in 2009, where he was mentored by Prof. Weihong Tan. His research interest during graduate study is to develop DNA-based molecular system for intelligent therapy. He received his Ph.D. in chemistry from the University of Florida in the fall of 2013.