## **A cascificity minimizing the basic steps of acquired immune response** A cascade reaction network mimicking the basic functional

Da Han<sup>1, 2,\$</sup>, Cuichen Wu<sup>1, 2,\$</sup>,, Mingxu You<sup>1,\$</sup>,, Tao Zhang<sup>2</sup>, Shuo Wan<sup>2</sup>, Tao Chen<sup>1,2</sup>, Liping Qiu<sup>1</sup>, Zheng Zheng<sup>2</sup>, Hao Liang<sup>1</sup>, and Weihong Tan<sup>1, 2,\*</sup>

1. Molecular Science and Biomedicine Laboratory, State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, College of Biology, Collaborative Innovation Center for Molecular Engineering for Theranostics, Hunan University, Changsha, 410082, China

<sup>2.</sup> Center for Research at Bio/nano Interface, Department of Chemistry and Department of Physiology and Functional Genomics, Health Cancer Center, UF Genetics Institute and McKnight Brain Institute, University of Florida, Gainesville, FL 32611-7200

\*Correspondence to: tan@chem.ufl.edu

\$: equal co-first authors.

**Supplementary Materials:**  Materials and Methods Detailed Discussion Figures S1-S13 Tables S1 References (*1-5*)

#### **1. Materials and Methods:**

**DNA synthesis**. DNA sequences were synthesized on an 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. HPLC purification was performed with a cleaned Alltech C18 column on a Varian Prostar HPLC machine. The collected DNA product was dried, and detritylation was processed by dissolving and incubating in 200 µL 80% acetic acid for 20 minutes. The detritylated DNA product was mixed with 400 µL ethanol and vacuum dried. The DNA products were quantified and stored in DNA water for subsequent experiments. The detailed sequence information is described in Table S1.

**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  $\mu$ M in 1×TAE-Mg buffer (20 mM Tris-Acetate-EDTA, pH 7.5, 12.5 mM  $Mg(Ac)_2$ ). Native PAGE gels (10%) in 1×TAE-Mg buffer were run at 100 V for 90 minutes at 4°C and stained with Gel Red stain solution (Biotium, Hayward, 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, Billerica, MA). Finally, the DNA duplex sequences were quantified by UV spectrometry and kept in buffer for future use.

#### **2. AIRS mimicry design rules**

We hypothesized the feasibility of designing a network of chemical reactions able to function in a manner similar to the vertebrate acquired immune system, using DNA and enzyme as simplified artificial analogs to mimic the activation and deployment of acquired immune response at its most fundamental level. This chemical reaction network, termed AIRS, couples DNA-enzyme cascade interactions with DNA strand displacement cascades in which a new product can be activated by the presence of some initiators, thus allowing many such reactions to be linked into a cascade and even form a complex network, in particular one such as adaptive immunity where biological events follow each other sequentially in a rolling fashion.

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 adaptive, or acquired, immune system is composed of both humoral and cell-mediated components that are called to action when immune tolerance is surpassed. More specifically, adaptive immunity involves the production of antibodies and immunological memory, but also cell-mediated responses, including the key steps of antigen presentation, antigen binding, co-stimulation and destruction of the foreign body (e.g.,

phagocytosis). A key function is antigen presentation which provides the mechanism for recognition and response. While B cells secrete antibodies, which bind to and label an antigen, T cells perform the job of attacking target cells. Both T and B cells have roles in immunological memory. We have divided this series of events into three steps, including Recognition and Tolerance, Immune Response, and Killing and Memory, under the control of four functional DNA components, including DNA duplexes **AM** (**A**ntigen Presenting Cell (APC) **M**imicry), **BM**  (**B** Cell **M**imicry), **PG** (**P**rimer **G**enerator, which generates primers for analog antibody) and single-stranded circular DNA **CT** (**C**ircular **T**emplate, which controls the sequence of analog antibody), as well as two enzymes (Phi29 DNA polymerase and SsPI restriction enzyme) able to autonomously and programmably respond to an incoming piece of single-stranded DNA pathogen input  $(P_0)$  taken from bacterial genome. When no  $P_0$  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 increased load of viral  $P_0$ , these functional domains are activated in a series of steps designed to mimic the three steps noted above.

## **Step 1: Recognition and Tolerance**

In AIRS, this step depends on the core sequence of pathogen input  $P_0$ , which is composed of two effective domains. The first domain, (**P**), is a genomic sequence taken from *B. anthracis* for pathogen mimicry. The second domain, (domain 2-3-4), is designed for systematic cascade reaction, i.e., DNA displacement reaction in steps 1 and 2. Accordingly, **AM**, **BM** and **PG,** as defined above, were all designed to interact with  $P_0$ . All DNA strands in the circuit consist of recognition domains (20-44 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. To differentiate the displacement reaction rates between  $P_0$  and  $AM$ , as well

as between  $P_0$  and **BM**, a 10-nt long toehold  $(2a*3*)$  was used in **AM**, and a 0-nt long toehold was designed in **BM**. As discussed in the previous kinetic model above, the predicted reaction rate differences between  $P_0$  and **AM** or  $P_0$  and **BM** are about 1000 x.<sup>1</sup>

#### **Step 2: Immune Response**

When immune tolerance has been surpassed, antigen-specific immune response is triggered to eliminate foreign agents. Many mechanisms act in a coordinated and sequential manner to recruit different cell populations, mainly B and T cells, and promote the production of antibodies, as well as proinflammatory cytokines and chemokines. For example, after antigens have been transported to lymph nodes, they are digested into smaller fragments for presentation by Antigen Presenting Cells (typically B cells), allowing for T cell recognition and the initiation of cytotoxic processes, such as the production of cytokines and antibodies.

First, it is enzymatic amplification (RCA) that controls the amount of antibody mimicry **P\*** generated. As indicated previously, an efficient isothermal enzyme-based amplification, termed rolling-circle amplification (RCA), provides us with an excellent way to specifically produce a large amount of desired DNA product in a short time and at a constant temperature, thus mimicking the generation of antibody in our network. To demonstrate that the concentration of RCA product depends on the amount of primer, different concentrations of ssDNA primer were incubated with 50 nM CT, 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 *B. anthracis* genomic sequence (**P)** as its loop. **MB-R** could be opened by the RCA products, i.e., **P\***, thereby exhibiting fluorescence restoration that represents different amounts of **P\***. Figure S8A shows the resultant fluorescence curve establishing the

quantitative relationship between **P\*** and released primer. This allowed the quantity of **P\*** generated from enzymatic amplification for hybridization with domain **P** in  $P_0$ ,  $P_1$  and  $P_2$  to be determined by the upstream reactions in steps 1 and 2.

Next, to design a more sensitive reporting system to indicate the amount of generated RCA product (**P\***), we applied the amplification circuit motif comprised of two DNA hairpins developed by Li et al.<sup>2</sup> and Jiang et al.<sup>3</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**), **A1** and **A2** can form a stable duplex without consuming **C**. As shown in Figure S9, **A1**, **A2** 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 **A2**. 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. The overall reaction is driven by a decrease in enthalpy resulting from the formation of  $A_{12}$  with a greater number of base pairs.

To apply this amplification circuit for detecting the amount of **antibody P\***, we encoded the sequence **P\*** into the catalyst **C** sequence and designed **A1** and **A2** correspondingly. Because of the high amplification efficiency of the circuit, this system can report the amount of **P\*** more sensitively compared to the molecular beacon-based method (**MB-R**). As shown in Figure S10, when the amount of pathogen is under the level of immune tolerance (threshold of 200 nM), it

can be observed that the fluorescence kinetics is similar to that without  $P_0$ . However, when another 150 nM  $P_0$  were added to the system after the first addition of 100 nM to mimic a second exposure to the same pathogen, the fluorescence restoration showed much faster kinetics behavior.

It was found that sequence design in this system could be affected by unspecific elongation of polymerase Phi29. Such unexpected elongation could represent a serious design flaw because it can seal all the toehold regions that initiate the strand displacement. Therefore, since Phi29 is a universal polymerase which can only recognize natural 3'-end of a DNA duplex, we first modified an artificial Inverted dT group at the 3'-end of each DNA component. As a consequence, DNA duplex could form a parallel linkage (3'-3') to impart duplex exonuclease resistance and thus prevent extension by polymerases in the absence of free 3' hydroxyl group to initiate synthesis.<sup>4, 5</sup> Importantly, however, the primer strand  $(12a)$  cannot be modified with Inverted dT group, as it is designed for polymerase recognition and elongation and should not be initially blocked. Therefore, to prevent unexpected elongation from the 3'-end of primer strand, which could result in sealing the toehold of **PG** duplex when incubated with Phi29, we designed 1 or 2 mismatch points between the primer (12a) and its complementary strand (1\*2a\*b\*) in **PG** at the 3'-end of primer (Fig. S8B). An experiment was carried out to demonstrate the ability of these mismatched points to prevent the undesired sealing of Phi29. Specifically, these mismatched **PG** duplexes labeled by FRET pairs (FAM-DABCYL) were incubated with 0.1 U/µl or 1 U/µl Phi29 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** were found to be sealed by Phi29, the results would show the displacement reaction between **PI** and **PG** to be negligible, leading to low fluorescence recovery. As shown in Figure

S8, 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 leakage in the system. Thus, 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.

### **Step 3: Selective killing effect**

The sequence of **CT** was designed by encoding two *Bacillus anthracis* genomic sequence fragments (**P**) inside. These two **P** sequences were separated by a linker composed of ten T nucleotides. Here, because of the encoding of *Bacillus anthracis* genomic sequence fragments (**P**) in CT, the generated RCA product ( $P^*$ ) is only able to recognize pathogen  $P_0$ , thus providing the first layer of systematic selectivity. Additionally, **P<sub>0</sub>-mis-P**\* mismatch duplex cannot form an effective recognition binding site with the restriction enzyme SsPI (AAT-ATT), thereby reducing the enzyme's digestion rate, which, in turn, results in much less cleavage product from input  $P_0$ -mis (Figure S11). Therefore, the design of this AIRS mimicry system possesses the character of antigen-specific immune response similar to that of the actual host immune defense system.



**Figure S1**. Kinetics experiments of the system with 50 nM **AM**, 50 nM **BM** and 50 nM **PG**. 30  $nM$  **P**<sub>0</sub> were added to the buffer separately at each time point.



**Figure S2**. PAGE (10% native gel) analysis of the reaction pathway in steps 1 and 2. Here, to fully separate DNA bands with similar molecular weight, the  $P_0$  sequence was modified by eliminating the *Bacillus anthracis* genomic sequence part (**P**: 28 bp). The products and their base numbers are indicated on the right.



**Figure S3.** Fluorescence studies of the reaction priority between steps 1 and 2 with  $P_0$ -SARS, as well as the catalytic effect of **TM** on the reaction of **BM** and **P<sub>0</sub>-SARS**. (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  $P_0$ -SARS. (**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>-SARS** were added to the buffer separately at each time point. Experiments were performed at 25°C in 50 mM Tris-HCl buffer containing  $10 \text{ mM } MgCl<sub>2</sub>$ .



**Figure S4**. Scheme of the primer-initiated RCA product cut by SsPI restriction enzyme and the 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  $P_0$ -SARS+2U/ $\mu$ L SsPI. Lane 2: RCA product + 1  $\mu$ M FAMlabeled  $P_0$  +2U/ $\mu$ L SsPI. Lane 3: FAM-labeled  $P_0$ .



**Figure S5**. Gel analysis by PAGE (8% denatured gel) for the entire AIRS 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 **CT**, 0.5 U/ µL Phi29 and 250 µM dNTP. Lane 1: FAMlabeled **P<sub>0</sub>**. Lane 2: M + 1  $\mu$ M FAM-labeled **P**<sub>0</sub>+ 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 + 50 nM **CT** + 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 S6**. Analysis by PAGE (8% denatured gel) of the preparation of circular template (**CT**). Lane1: After ligation. Lane 2: Before ligation.



**Figure S7**. The proposed entropy-driven catalytic pathway.<sup>1</sup> **TM** first binds with **BM** through a 4-nt toehold region (domain 5) and forms intermediate  $I_1$  and ssDNA  $W_1$ .  $P_0$  can bind with  $I_1$  by a newly formed 4-nt toehold (domain 3) and displace **TM** and ssDNA **PI**. The hybridization reaction is driven forward thermodynamically by the entropic gain of the liberated molecules.



**Figure S8**. (A) Scheme showing the primer-assisted RCA reaction and plot of fluorescence restoration versus different concentrations of primer. (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 50 mM Tris-HCl buffer containing 10 mM  $MgCl<sub>2</sub>$ , 10mM  $(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>$  and 4mM dithiothreitol (DTT).



**Figure S9**. (A). Scheme of detailed Catalytic Hairpin Assembly (CHA) reaction of DNA hairpins **A1** and **A2** catalyzed by **C** sequence. Different domains are labeled with different colors. All **x** domains are complementary to **x\***. (B) Scheme of the Catalytic Hairpin Assembly (CHA) circuit for reporting the amount of **P\***. The circuit involves two individual steps. In the catalytic step,  $P^*$  catalyzes DNA hairpins  $A_1$  and  $A_2$  to form duplex  $A_{12}$ . In the fluorescent reporting step,  $A_{12}$  can open duplex  $R_{12}$  and displace quencher-labeled single-strand  $R_2$  to form  $A_{12}-R_1$ . Subsequently, FAM-labeled  $\mathbf{R}_1$  can be released to indicate the amount of  $\mathbf{P}^*$ .



**Figure S10**. Fluorescence kinetics of Catalytic Hairpin Assembly (CHA) containing **A1**, **A2** and **R12** to detect the amount of **P\***. 1) CHA only: A mixture of 200 nM **A1**, 200 nM **A2** and 250 nM **R12**. 2) CHA+AIRS mimicry: 1+200 nM **AM**, 200 nM **BM**, 200 nM **PG**, 50 nM **CT**, 0.5 U/ µL Phi29 and 250  $\mu$ M dNTP. 3) 2+100 nM **P**<sub>0</sub>. 4) 2+100 nM+150 nM **P**<sub>0</sub>.

# DOI: 10.1038/NCHEM.2325 SUPPLEMENTARY INFORMATION



**Figure S11**. Analysis of the reactions with  $P_0$ -mis in the entire system by agarose gel (1%). M = mixture of 200 nM **AM**, 200 nM **BM**, 200 nM **PG**, 50 nM **CT**, 0.5 U/µL Phi29 and 250 µM dNTP. Incubation time = 1 hour. Lane 1:  $M + 0$  nM FAM-labeled  $P_0$ -mis; lane 2:  $M + 100$ nM FAM-labeled **P<sub>0</sub>-mis**; lane 3: M + 500nM FAM-labeled **P<sub>0</sub>-mis**; lane 4: M + 100nM FAMlabeled  $P_0$ -mis + 2U/µL SsPI; lane 5: M + 500nM FAM-labeled  $P_0$ -mis + 2U/µL SsPI; L: 1kb ladder.



Figure S12. Agarose gel analysis of system specificity (1%). The image was taken under FITC channel. FITC imaging channel: EX=488nm, EM=512nm. EB imaging channel: EX=420 nm, EM=605 nm.



**Figure S13**. Analysis of AIRS analog capacity by agarose gel (0.8%). M = 200 nM **AM**, 200 nM **BM**, 200 nM **PG**, 50 nM **CT**, 0.5 U/µL Phi29 and 250 µM dNTP in the NEB buffer. Lane 1: M  $+ 200$  nM FAM-labeled  $P_0$ ; lane 2: M + 1  $\mu$ M FAM-labeled  $P_0$ ; lane 3: M + 2  $\mu$ M FAM-labeled **P**<sub>0</sub>; lane 4:  $M + 5 \mu M$  FAM-labeled **P**<sub>0</sub>; lane 5: **P**<sub>0</sub> only.



**Figure S14**. Fluorescence validation of entire AIRS mimicry network with two pathogen inputs (**P0** and **P0-SARS**) present simultaneously. 300nM **AM**, 300nM **BM**, 300nM **PG**, 50nM **CT-P0**, 50nM **CT-P-SARS**, 0.5U/ul Phi29 and 250uM dNTP were mixed with 500nM FAM-labeled **MB-R and MB-R-SARS.** Different amounts of  $P_0$  and  $P_0$ -SARS were added to the mixture, and the fluorescence was measured, following 30 min incubation. Experiments were performed at  $25^{\circ}$ C in 50 mM Tris-HCl buffer containing 10 mM MgCl<sub>2</sub>.

## Table S1: Sequences of AIRS mimicry





### **References:**

- 1. Zhang, D. Y., Turberfield, A. J., Yurke, B. & Winfree, E. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318**, 1121-1125 (2007).
- 2. Li, B.L., Ellington, A. D. & Chen, X. Rational, Modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Res.* **39,** e110 (2011).
- 3. Jiang, Y.S., Li, B., Milligan, J.N., Bhadra, S. & Ellington, A.D. Real-time detection of isothermal amplification reactions with thermostable catalytic hairpin assembly. *J. Am. Chem. Soc.* **135**, 7430-7433 (2013).
- 4. Ortigao, J.F. 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).
- 5. Vandesande, J.H. et al. Parallel Stranded DNA. *Science* **241**, 551-557 (1988)