# Constructing Smart Protocells with Built-In DNA Computational Core to Eliminate Exogenous Challenge

**Yifan Lyu**[†,‡,§], **Cuichen Wu**[†,§], **Charles Heinke**[§], **Da Han**[*,‡,§], **Ren Cai**[†,§], **I-Ting Teng**[§], **Yuan Liu**[†,§], **Hui Liu**[†], **Xiaobing Zhang**[†], **Qiaoling Liu**[†], and **Weihong Tan**[*,†,‡,§]

[†]Molecular Science and Biomedicine Laboratory, State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, College of Life sciences, Aptamer Engineering Center of Hunan Province, Hunan University, Changsha, Hunan 410082, China

[‡]Institute of Molecular Medicine, Renji Hospital, School of Medicine and College of Chemistry and Chemical Engineering, Shanghai Jiao Tong University, Shanghai 200240, PR China

[§]Department of Chemistry and Department of Physiology and Functional Genomics, Center for Research at the Bio/Nano Interface, UF Health Cancer Center, UF Genetics Institute, McKnight Brain Institute, University of Florida, Gainesville, Florida 32611-7200, United States

## Abstract

A DNA reaction network is like a biological algorithm that can respond to "molecular input signals", such as biological molecules, while the artificial cell is like a microrobot whose function is powered by the encapsulated DNA reaction network. In this work, we describe the feasibility of using a DNA reaction network as the computational core of a protocell, which will perform an artificial immune response in a concise way to eliminate a mimicked pathogenic challenge. Such a DNA reaction network (RN)-powered protocell can realize the connection of logical computation and biological recognition due to the natural programmability and biological properties of DNA. Thus, the biological input molecules can be easily involved in the molecular computation and the computation process can be spatially isolated and protected by artificial bilayer membrane. We believe the strategy proposed in the current paper, i.e., using DNA RN to power artificial cells, will lay the groundwork for understanding the basic design principles of DNA algorithm-based nanodevices which will, in turn, inspire the construction of artificial cells, or protocells, that will find a place in future biomedical research.
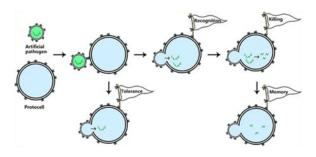
## Graphical Abstract

[*]**Corresponding Authors** W.T. Fax: (+1) 352-846-2410, tan@chem.ufl.edu, D.T., handa806@gmail.com.

## INTRODUCTION

Protocells, as the most simplified prototype of artificial cell, have previously been exploited to mimic some basic cellular behaviors, such as selective membrane permeability, DNA-directed protein translation,[1,2] cell–cell communication, cellular ingestion and cell division.[3–8] Such kinds of protocells, though powerful as research tools of basic cellular properties, sometimes are not very "smart" from the perspective of bioengineering. In nature, a series of temporospatially ordered chemical reactions, also known as reaction networks (RNs), combine to regulate different biofunctions that guarantee the survival of all organisms.[9] These intricate and sophisticated natural RNs transmit information and respond to external stimuli in a way that ensures organismal homeostasis.[10] Accordingly, it is both challenging and attractive to engineer protocells encapsulating an artificial RN as computational core which is able to perform a programmed function.[11–13] To date, however, studies of a protocell still mainly focuses on properties of bilayers[6,14] or on-membrane structures,[15,16] while protocells containing artificial RNs for logical computation has never been attempted.

Benefiting from the specificity and predictability of Watson–Crick base pairing and mature chemical synthesis, DNA molecules are regarded as natural building blocks and versatile biomaterials for nanoscale engineering.[17] With the development of DNA nanotechnology, dynamic DNA RNs have been designed with numerous robust functions and used to run some fundamental algorithms,[18,19] demonstrating the feasibility of linking biomolecules to algorithm-based nanostructures. Although no man-made chemical RNs, including DNA RNs, have been able to match electronic circuitry such as that found in silicon chips, DNA RNs running in biological environments can directly interact with biological targets, such as other DNA strands,[20] macrobiomolecules,[21] or even small organisms,[22] in a natural way. However, DNA RNs thus far reported are typically run in huge containers like tubes, along one- or two-dimensional tracks or have to be delivered into/onto cells for amplified detection or signal transduction.[22–25] Some of these diffusible DNA RNs, though compartmentalized in a comparatively stable environment, may still be easily disturbed by other macromolecules or inactivated due to the extremely diluted concentration when working in complex electrolyte-rich environments like body fluid.[26] We previously designed a cascade RN using DNA and enzymes, as simplified artificial analogues, to mimic the activation and deployment of the vertebrate adaptive immune response (AIS)[27] at its most basic level. This network, termed "artificial immune response simulator" (AIRS), has been proven to work effectively in a test tube against a pathogenic DNA sequence taken from the genome of severe acute respiratory syndrome coronavirus (SARS coronavirus, or SARS-CoV), which

has a lipid-containing envelope. It should be noted that SARS-CoV attaches to the host cell through binding between the receptor binding domain (RBD) in the S1 subunit of S (Spike) protein to the cellular receptor angiotensin-converting enzyme 2 (ACE2)[28] and enters the cell via endocytosis.[29,30] Herein, by building AIRS inside of a protocell, we further built a cell-like robot whose biofunctions are programmed by encapsulated DNA logic RN, thereby making AIRS the actual computational core of the protocell (Figure 1a). Such a DNA RN-powered protocell can realize the connection of molecule-dependent logical computation and biological recognition due to the natural programmability and biological properties of DNA, so biological input molecules can be easily involved in the molecular computation. Since our DNA RN is spatially isolated and protected by artificial bilayer membrane, a locally high concentration of computing molecules is permitted and a relatively reliable molecular computation process is allowed, just like the functions of membranaceous structures of cells and organelles.

By constructing an artificial pathogen, mimicked host immune response can happen in a way defined by AIRS but within an artificially constructed intracellular microenvironment (Figure 1b) to eliminate the artificial exogenous pathogenic challenge. Delivery of a small amount of pathogen DNA into the protocell triggers the first step of built-in AIRS: recognition and tolerance. As long as the amount of pathogen DNA is below the threshold of immune tolerance, the system is inactive. However, in the presence of excessive pathogen DNA, rolling circle amplification (RCA)-based[31,32] artificial immune response is activated, involving the production of antibody mimicry, immunological memory, and the destruction of the foreign DNA (see Figure 1c). To the best of our knowledge, this is the first experiment of its kind in which a DNA logic RN has been built as the computational core of a protocell for programmable biomimicry.

## EXPERIMENTAL SECTION

### Computational Simulation Using Visual DSD.

The DNA networks of steps 1 and 2 in AIRS were first validated using visual DSD, an open access software which allows design and analysis of DNA strand displacement reactions.[33] All the simulations were run in detail mode. Programs used in visual DSD and simulation results are shown in Figures S1–S3.

### DNA Synthesis.

DNA synthesis reagents were purchased from Glen Research (Sterling, VA). 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 mL of AMA (ammonium hydroxide/methylamine 50:50) for 17 h at 40 °C in water bath. The cleaved DNA product was transferred to a 15 mL centrifuge tube and mixed with 200 $\mu$L of 3.0 M NaCl and 5.0 mL of ethanol, after which the sample was placed into a freezer at –20 °C for ethanol precipitation. Afterward, the DNA product was spun at 4000 rpm at 3 °C for 20 min. The supernatant was removed, and the precipitated DNA product was dissolved in 400 $\mu$L of 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. An aqueous solution of 0.1 M triethylamine acetate (pH 6.5) was used as HPLC buffer A, and HPLC-grade acetonitrile from Oceanpak (Sweden) was used as HPLC buffer B. The collected DNA product was dried, and detritylation was processed by dissolving and incubating in 200 $\mu$L of 80% acetic acid for 20 min. The detritylated DNA product was mixed with 400 $\mu$L of 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 RN Purification.

Native PAGE was applied to purify AM and BM (BM-FQ) duplex strands to remove excess strands and avoid undesired system leakage caused by sequence defects. The ssDNA components of AM and BM (BM-FQ) were annealed at concentrations of around 50 $\mu$M in $1\times$ TAE/Mg$^{2+}$ buffer (40 mM Tris-acetic acid, 1 mM EDTA, and 12.5 mM magnesium acetate, pH 8.0). Native PAGE gels (12%) in $1\times$ TAE/Mg$^{2+}$ buffer were run at 120 V for 60 min at 4 °C and stained with Stainsall stain solution (Sigma-Aldrich). Only the sharp bands were cut from the gels, chopped into small pieces, and soaked in $1\times$ TAE/Mg$^{2+}$ buffer for 24 h. After extracting most of the DNA 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. The extinction coefficients for double-stranded complexes were calculated using the following equation:

$$\varepsilon = \varepsilon(\text{top strand}) + \varepsilon(\text{bottom strand}) - 3200N_{\text{AT}} - 2000N_{\text{GC}}$$

where $N_{\text{AT}}$ and $N_{\text{GC}}$ are the numbers of AT pairs and GC pairs in the double-stranded domain, respectively.[19,34]

### Preparation of DNA Circular Template.

A 200 $\mu$L sample of 10 $\mu$M CT (with a phosphate group on its 5′ end) and 20 $\mu$M AI-T (with an inverted T base on the 3′ end) were annealed in $1\times$ T4 DNA ligase reaction buffer containing 50 mM Tris-HCl, 10 mM MgCl$_2$, 1 mM ATP, 10 mM DTT, and pH 7.5. Afterward, 10 $\mu$L of T4 DNA ligase was mixed with the solution. The mixture was incubated at room temperature overnight. Then, 200 U of Exonuclease I and 2000 U of Exonuclease III (New England Biotech) were added to the mixture and incubated at 37 °C for 1 h. The enzyme was denatured by heating the solution to 90 °C for 20 min. Afterward, the CT product was purified by denatured PAGE and HPLC and then desalted with NAP-5 columns. Finally, ligated CT was quantified by UV spectrometry and kept in buffer for future use.

### Electrophoresis Analysis of the System.

A 10 $\mu$L system with purified AM (500 nM), BM (500 nM), CT (50 nM), Phi29 (0.5 U $\mu$L$^{-1}$) and dNTP (250 $\mu$M) was chosen for this experiment. For the pathogen-digestion experiment, the concentration of restriction enzyme *Ssp*I was set at 2 U $\mu$L$^{-1}$. The gel was

run at 4 °C for 1 h at 120 V. Twelve % SDS-polyacrylamide gel with 6 M urea was used to analyze the protocells. A 100 $\mu$L sample of protocells was dried under reduced pressure and redissolved with 10 $\mu$L of 1× TBE buffer (90 mM Tris-borate, 2 mM EDTA) with 1‰ SDS. Samples were supplemented with 2 $\mu$L of 5× DNA ladder and 10 uL of saturated urea solution before loading into each well. TBE buffer (1× TBE) with 1‰ SDS was used as electrophoresis running buffer, and the gel was run at 110 V for 1 h. The buffer temperature was controlled to maintain the samples at 4 °C throughout the run. The gel was then stained with SYBR Gold nucleic acid gel stain (Sigma-Aldrich) stain solution and imaged by a Typhoon gel reader.

### Validation of Signal Transduction in Each Step and the Entire System by Fluorescence.

When the reaction priority between steps 1 and 2 was tested, purified AM (60 nM) and BM/BM-FQ (100 nM) were mixed in 1× TAE/Mg$^{2+}$ buffer to a total volume of 100 $\mu$L, and the fluorescence was monitored for different concentrations of strand P. When generated antibody mimicry was tested by fluorescence, purified AM (100 nM), BM (100 nM), CT (50 nM), and antibody-MB (500 nM) were incubated in 1× Phi 29 DNA polymerase buffer. The fluorescence kinetics was monitored at different time points after adding different amounts of strand P. All fluorescence measurements were performed on a Fluoromax-4 spectrofluorometer from Horiba with temperature controller, using a quartz fluorescence cell with an optical path length of 1.0 cm. For spectrofluorimetry studies, the excitation was recorded at 488 nm with recording emission range of 500–600 nm. For fluorescence kinetics studies, the excitation was recorded at 488 nm, and the emission was recorded at 520 nm. Unless otherwise specified, all excitation and emission bandwidths were set at 5 nm. Prior to each experiment, all cuvettes were washed with 70% ethanol and distilled water.

### Preparation of Artificial Pathogen and Protocell.

DOPC, DOPG, EDOPC, and cholesterol were purchased from Avanti Lipids Inc. (Alabaster, AL, USA). Artificial pathogens and protocells were prepared by the hydration and extrusion method.[35] For the preparation of artificial pathogens, stock chloroform solutions of DOPC and EDOPC were mixed in a glass trial bottle to a final molar ratio of DOPC/EDOPC = 9:1. The mixture was dried to a thin lipid film under a stream of nitrogen and further dried under reduced pressure for another 5 h to remove traces of solvent. The dry lipid film was hydrated by the addition of 2 $\mu$M of pathogen DNA in phi 29 polymerase buffer solution and vortexed for 5 s, resulting in the formation of a small vesicle dispersion. The dispersion was incubated for at least 1 h at room temperature to stabilize the vesicles and then extruded through a 200 nm polycarbonate membrane (Whatman) to homogenize vesicles. Then, 500 nM of SP in Phi 29 DNA polymerase buffer solution was added to the dispersion. The mixture was incubated for more than 1 h at room temperature to allow nearly all the cholesterol-modified DNA to assemble onto the vesicle surface. Unloaded DNA was removed by cl-2b column. For protocells, stock chloroform solutions of DOPC and DOPG were mixed in a glass trial bottle to a final molar ratio of DOPC/DOPG = 1:4. The mixture was dried to a thin lipid film under a stream of nitrogen and further dried under reduced pressure for another 5 h to remove traces of solvent. The dry lipid film was hydrated by a solution of 60 nM AM, 100 nM BM, 50 nM CT, 250 $\mu$M dNTP, 0.5 U/$\mu$L Phi 29 DNA polymerase, and 2 U/$\mu$L Ssp I endonuclease in Phi 29 DNA polymerase reaction buffer and vortexed for 5 s, resulting in

the formation of a vesicle dispersion. The dispersion was incubated for at least 1 h at room temperature to stabilize the vesicles and then extruded through a 2 $\mu$m polycarbonate membrane (Whatman) to homogenize vesicles. Then, 500 nM of SC in Phi 29 DNA polymerase buffer solution were added to the dispersion. The mixture was incubated for more than 1 h at room temperature to allow nearly all the cholesterol-modified DNA to assemble onto the vesicle surface. Unloaded DNA was removed by cl-2b column. The derived count rate acquired from Malvern Zetasizer was used for relative quantification of artificial pathogens and protocells.[36,37]

### Pathogen Infection Mimicry and Confocal Imaging of Protocell.

Pathogen infection mimicry was achieved by mixing different volumes of artificial pathogens with protocells in a tube. After incubation for several hours at room temperature, 200 $\mu$L of infected protocells were subjected to confocal microscopy imaging using a X81 laser confocal microscope (Olympus) in sequencing scan mode.

### TEM Imaging of Artificial Pathogen.

A negative-staining method was used for TEM imaging of artificial pathogen. First, a carbon-coated copper grid was electrostatically discharged and then put face down on the artificial pathogen stock solution for about 1 min. Then, the copper grid was moved face downward onto the negative-staining solution (2% uranyl acetate water solution (w/v)) for another 3–5 min. Then the sample was placed face up onto a piece of filter paper and dried at room temperature. Afterward, the sample was inserted into the F-2010 transmission electron microscope and imaged at a working voltage of 100 kV. The size of artificial pathogen evaluated by TEM imaging was slightly smaller compared with DLS results because the membrane had shrunk during the drying process, which is normal for vesicle structures.

## RESULTS AND DISCUSSION

This molecular computing process in a protocell begins with a pathogenic infection mimicked by DNA migration and electrostatic interaction-mediated membrane fusion between artificial pathogen and protocell. As noted above, with the help of the receptor-binding domain (RBD), a small fragment in the S1 domain of spike (S) protein, previous studies have demonstrated that SARS-CoV can infect the target cell through endocytosis after initial attachment.[38] Since the sequence of pathogen DNA used in this work was also taken from the genome of SARS-CoV, we chose a simplified receptor-mediated membrane fusion as our artificial infection model by constructing artificial pathogen with lipid bilayer as we will see in this paper. Once invading the protocell, a piece of pathogen DNA is involved into the computational core and triggers the built-in AIRS, ending with its enzymatic, i.e., phagocytic, digestion.

### Step-by-Step Working Principle of DNA Computational Core in Protocell.

The triggered AIRS for protocell after pathogen invasion is shown in Figure 1b,c with three steps, including recognition and tolerance, immune response, and killing and memory. For a comprehensive description, a step-by-step working principle of AIRS follows.

**Step 1.—**As soon as a small amount of pathogen DNA (P) is delivered inside the protocell, as noted above, antigen-presenting cell mimicry (AM) first recognizes pathogen DNA and labels it as pathogen with label (PL, Label strand). In this case, only step 1 is activated because the amount of pathogen DNA does not exceed the level of immune tolerance. This event is followed by the release of T cell mimicry (TM) which will act as a catalyst as we will see in step 2. All these operations, as well as most operations in the following two steps, are controlled by toehold-mediated strand displacement reactions to guarantee the highly efficient and temporospatially ordered information transfer of AIRS.

**Step 2.—**In AIS, once invaded pathogen has crossed the boundary of immune tolerance, both humoral and cell-mediated immune responses are called to action. Helper T cells (Th) produce chemicals that trigger B cells to develop into plasma cells, while others stimulate killer T cells to target and kill cells that become infected. Plasma cells are antibody-manufacturing cells derived from B lymphocytes, following their activation by an antigen. These B cells express B cell receptors (BCRs) on their cell membranes. BCRs allow the B cell to bind a specific antigen, against which it will initiate an antibody response. In AIRS, these operations are controlled in concise and effective ways. First, as noted in step 1 above, TM is released and acts as a catalyst (Figure 2a) that increases the rate of immune response against excessive pathogen DNA, thus activating B-cell mimicry (BM) which results in the release of an antibody initiator strand (AI). When pathogen DNA reacts with BM in this step, they are also labeled as PL. Second, the generated antibody mimicry (RCA product of a circular template (CT)) is a concatemer containing tens to hundreds of tandem repeats that are designed to be complementary to pathogen DNA to efficiently follow pathogen DNA capture.

**Step 3.—**In contrast to the specific binding between antibody and antigen in antibody-mediated immune response, antibody mimicry in AIRS captures both free and labeled pathogen DNA (both P and PL) via hybridization. The antibody mimicry-pathogen DNA complex is cut into pieces by *Ssp*I, a restriction enzyme that cuts DNA at a specified site. Activated TM is not consumed and remains in the system once it is released in step1. Consequently, succeeding pathogen DNAs will not need to be recognized in the first step and can directly trigger the immune response in steps 2 and 3. This is considered the memory property of AIRS, mimicking, as noted above, what is known as immunological memory in the AIS.

## Step-by-Step Construction of DNA Computational Core.

Steps 1 and 2 in AIRS are totally DNA strand-displacement-based RNs so that Visual DSD software[33] was implemented to validate the function of each step *in silico*, which showed the reliability of our system because each step was simulated to perform as accurately as the design (Figures S1–S3). These two steps were first constructed in buffer solution. In order to monitor steps 1 and 2, we modified the BM complex to BM-FQ by adding a quencher (dabcyl) and a fluorophore (FITC) (Figure 2b). In the initial state, the fluorophore is quenched as a result of hybridization. When a small amount of pathogen DNA triggers step 1, AM functions as a threshold gate, and no fluorescence is observed, as shown in Figure 2d,e. However, upon the introduction of excessive pathogen DNA, step 2 is triggered, and

fluorophore-modified AI is replaced by pathogen DNA, leading to fluorescence recovery (Figure 2d,e). This process was also simulated using Visual DSD software and the simulation result matched perfectly with the experimental result (Figure 2d, dashed line). Electrophoresis results of step 1 and 2 are shown in Figure S4. The catalytic function of TM was illuminated by visual DSD simulation as well and then validated by experimentally monitoring released AI (Figure S5). Steps 2 and 3 were then confirmed with electrophoresis. Figure S6 shows that released AI hybridizes with a predesigned circular template and triggers RCA in the presence of Φ29 DNA polymerase, an enzyme from the bacteriophage Φ29. In the artificial system, the antibody production generated with RCA captures pathogen DNA via hybridization, and finally, with the help of endonuclease *Ssp*I, pathogen DNA is cut into small pieces and loses its pathogenic property (Figure S7), as designed in Figure 1b.

All three steps were subsequently combined into a single electrophoresis gel. As shown in Figure S8, compared with AIRS with no pathogen DNA (lane 6) or small amount of pathogen DNA (lane 5), only in the presence of pathogen DNA in an amount that exceeds the defined threshold (lane 4), all three steps are triggered. Antibody mimicry is generated and hybridizes with Pathogen DNA, followed by *Ssp*I-mediated digestion. We also designed a molecular beacon (Antibody-MB) to probe the generated antibody mimicry in real time (Figure 2c). We chose a domain in antibody mimicry as the target sequence and used a complementary sequence as the loop domain of Antibody-MB. Once antibody mimicry is generated during the active mode of AIRS, Antibody-MB is opened by hybridizing with the target domain on antibody mimicry. Real-time detection within 1 h is plotted in Figure S9a, which shows increasing fluorescence intensity after adding pathogen DNA. Data with longer reaction times up to 24 h were also measured and are shown in Figure S9b.

### Characterization of Artificial Pathogen and Protocell.

After confirming the reliability of AIRS for our protocell, we next built the artificial pathogen and protocell (Figures 1a and S10). The artificial pathogen consists of (1) a bilayer membrane structure composed of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) and 1,2-dioleoyl-*sn*-glycero-3-ethyl-phosphocholine (EDOPC), (2) a cholesterol-modified double-stranded DNA (dsDNA) inserted on the surface via hydrophobic forces between the cholesterol molecule and lipid monomer as SNARE protein[39] mimicry on pathogen (SP), and (3) a loaded pathogen DNA (a piece of SARS DNA; see Table S1) (Figure 3a). SNARE protein, an acronym derived from **SNA**P (Soluble NSF Attachment Protein) **RE**ceptor, plays a key role in numerous naturally existing membrane fusion events by generating the driving force needed to overcome repulsive electrostatic forces between the attached membranes. Here we used DNA strand migration to mimic the mechanism of SNARE protein (Figure 3e). The artificial pathogen has a diameter of less than 200 nm. The negatively stained TEM image showing the artificial pathogen's membrane structure is depicted in Figures 3b and S11.[40] Size-distribution, as determined by DLS, is shown in Figure 3c. The protocell was composed of (1) a larger bilayer membrane structure (less than 2 $\mu$m) made of 1,2-dioleoyl-*sn*-glycero-3-phospho-(1′-rac-glycerol) (DOPG) and DOPC, (2) cholesterol-modified dsDNA inserted on the surface as SNARE protein mimicry on cell (SC), and (3) encapsulated AIRS (Figure 3d). Figure 3f shows a single protocell with SC inserted on the

membrane (FITC labeled SC, marked as SC-F. See the structure shown in the middle.) By monitoring the fluorescent intensity distribution across a protocell, distinguishable signal can be confirmed on the membrane, indicating that cholesterol-modified dsDNA is effectively inserted on the membrane (Figure 3f, bottom). Lipid-sensitive dye Nile red, which fluoresces only in a hydrophobic environment, was also used to stain the protocell (Figure 3g, top and middle) and directly profile the membrane structure. Due to the amphiphilic property of the bilayer structure, red fluorescence only comes from the membrane of a protocell (Figure 3g, bottom). Figure 3h is the 3D reconstructed fluorescence image of a single protocell, showing the integral and stable chamber of a protocell.

### Encapsulation Efficiency.

Payloads were encapsulated in artificial pathogens and protocells during the hydration step.[41] A DNA model was first used to study encapsulation efficiency (Figures S12 and S13) of both artificial pathogen and protocell by encapsulating an FITC-labeled single-stranded DNA (ssDNA). Fluorescence images and flow cytometry results of protocell loaded with FITC-modified encapsulation probe EP-F are shown in Figure 4a–d. As a contrast, TAMRA-modified dsDNA SC-T was used to stain the membrane of protocell when performing confocal imaging. Confocal imaging in Figure 4a shows that EP-F was encapsulated into the chamber of a protocell. By analyzing the spatial distribution of signals in different channels, it is clear that the green fluorescence of FITC comes from inside of a protocell while red fluorescence of TAMRA comes from the membrane (Figure 4b), indicating effective encapsulation of payloads into protocells as shown in Figure 4c. The shift in FITC channel of flow cytometry also indicates the same result (Figure 4d). Rhodamine-attached BSA was then used as a protein model to confirm the encapsulation of protein into protocells due to its similar molecular weight with Φ29 (Figure S14). Artificial pathogens and protocells were then purified by size exclusion chromatography to remove unloaded DNA or enzymes (Figures S13, S15, and S16).[42,43]

### Infectivity and Pathogen DNA Delivery.

Infectivity of DNA pathogen upon encountering the protocell is driven by DNA migration and electrostatic interaction between membranes of artificial pathogen and protocell (Figure 3e).[44–49] Strand migration between SP and SC was demonstrated by electrophoresis (Figure S17a,c). dsDNA SL and SS were generated as the result of the migration reaction. With the help of double cholesterol modifications, SP and SC could stably insert into membranes of the artificial pathogen and protocell via hydrophobic assembly. Once artificial pathogen and protocell come into contact, their membranes adhere by the action of complementary toeholds of SP and SC. Then, the toehold-mediated migration between SP and SC and the opposite charge between membranes of artificial cell and protocell generates a force to pull the membranes off and mediate a fusion. In order to characterize the fused membrane, TAMRA-modified SP (SP-T) and FITC-modified SC (SC-F) were used, such that generated SS-FT would give a FRET signal in the TAMRA channel when excited at 488 nm (Figures 4e and S17b). This imaging result is shown in Figure 4f,g with faded green fluorescence in FITC channel and enhanced red fluorescence in TAMRA channel. The same were also confirmed by flow cytometry (Figure 4h). The insertion of SC-F onto protocell surface led to a forward shift in FITC channel (top left and top right in Figure 4h). After fusion with SP-T

modified artificial pathogen, a back shift in the FITC channel and a forward shift in the TAMRA channel can be observed when excited at 488 nm (top right and bottom right in Figure 4h). The hydrophobic dye molecule Nile red was also used to monitor the infection process by staining the artificial pathogen membrane. Upon attachment and fusion of artificial pathogen, Nile red was transferred into the membranes of protocells (Figure S18). Collectively, all these results indicated successful infection mimicry between artificial pathogen and protocell.

In order to further demonstrate that pathogen DNA had been delivered into the protocell as the result of pathogen infection, we next encapsulated EP-F in protocells and dabcyl-modified complementary strand cEP-Q in artificial pathogens (Figure S19a,b). Double-stranded SP and SC with no fluorophore modification were individually inserted into the membranes of artificial pathogen and protocells. As shown in Figure S19d, the protocell gave an obvious shift after being loaded with EP-F strands, indicating the encapsulation of DNA strands. After incubating with artificial pathogen containing cEP-Q, the infected protocells showed only relatively weak fluorescence intensity, which was regarded as a direct evidence of pathogen DNA delivery (Figure S19f). In contrast, when mixing the unencapsulated cEP-Q with EP-F loaded protocells, no back shift was observed (Figure S19e). The real-time fluorescence kinetics also indicated the same result (Figure S19c).

## AIRS in Protocell.

In order to ensure that our DNA computational core would function normally in micrometer-sized protocells, we first encapsulated all the reagents for AIRS, including pathogen DNA and Antibody-MB, in protocells. The mixture was first incubated on ice to prevent the simulator from activating before encapsulation. As shown in Figure S20, enhanced fluorescence intensity in the FITC channel was observed in flow cytometry peaks, demonstrating that AIRS works well, even in a cell-sized container.

We then used the artificial pathogen to infect the protocell through membrane fusion. Real-time monitoring of molecular computing process inside a protocell cell was accomplished through the addition of Antibody-MB to the AIRS mixture and loading into the protocells (Figure 5a). The generation of antibody mimicry was monitored by recording the incremental fluorescence emission of FITC (Figure 5b). Protocells (blue line) or protocells infected by small numbers of artificial pathogens (orange line), showed relatively low fluorescence intensity when compared with protocells infected with large numbers of artificial pathogens (red line). Because the fluorescent intensity is directly corelated with the generation of antibody mimicry, this result demonstrated the antigen-presentation and antibody-generation ability of the built-in AIRS upon challenge by an artificial foreign body. Confocal fluorescence imaging at different time point (2, 6, and 10 h) was also performed after infection mimicry to monitor the generation process of antibody mimicry (Figure 5c–e). Upon the introduction of an excessive amount of pathogen DNA delivered to protocells via membrane fusion, the threshold defined by the computational core was overwhelmed. The internalized pathogen DNA triggered all three steps of encapsulated AIRS, and antibody mimicry was generated as a result of artificial immune response. The antibody mimicry then hybridized with Antibody-MB and gave a green signal. Distinguishable fluorescence

enhancement was observed by analyzing each protocell during reaction times, indicating the continuous production of antibody mimicry. In contrast, without the addition of artificial pathogen, or in the presence of a small number of artificial pathogens, the protocells remained inactive with a very weak fluorescence signal (Figure 5f). Confocal images as well as corresponding fluorescent intensity distributions are shown in Figure 6a–d with obvious difference in FITC channel because only protocells infected with large number of artificial pathogens are activated. A *z*-axis scanning and 3D reconstruction of a triggered protocell are shown in Figure 6e, indicating that the green fluorescence came from opened Antibody-MBs inside a protocell and antibody mimicry was effectively generated as RCA products after the computation of DNA logic RN.

Antibody-MB allowed us to observe the generation of antibody mimicry, but the digestion of pathogen DNA could not be studied using confocal microscopy or flow cytometry. Therefore, we turned to denatured PAGE to analyze the results of triggered AIRS. We added 1% SDS and urea to the gel to individually denature membrane structure and DNA hybridization. According to the PAGE results shown in Figure 6e,f, a DNA band of about 50 nt was observed in the gel, and the length accurately matched the designed length of DNA fragment (49 nt) after digestion, indicating that part of pathogen DNA had been cut into smaller inert fragments, as predicted, providing evidence of the successful construction of protocell with built-in AIRS and the feasible molecular computing ability of DNA RN inside of a microsized protocell.

## CONCLUSION

Cells, as the basic functional units in a living organism, are composed of various sophisticated signal pathways that are isolated by a series of relatively stable membrane structures. Although the creation of a fully functional cell from scratch is a challenge, it is still valuable to construct artificial cells which mimic some of the cellular functions. DNA artificial RNs allow customization of special algorithms with logical operations and scaling abilities. In contrast with silicon-based logic circuits, DNA RNs can interact directly with biomolecules in living systems. Furthermore, the electrolyte-rich environment and the micrometer size of a cell inhibit intracellular application of silicon chips.

In this work, we described the feasibility of using a DNA RN as the computational core of a micron-sized protocell, which can perform an artificial immune response when there is an exogenous challenge. In order to validate the functionality of such protocell, an artificial pathogen DNA was also constructed and encapsulated therein. The *in silico* simulation using visual DSD and *in vitro* monitoring demonstrated the reliability of the DNA RN-based computational core. We also mimicked the process of pathogen infection by performing DNA migration/electrostatic-interaction-mediated membrane fusion, so pathogen DNA could be delivered into the protocell and trigger the encapsulated computational core. The designed functions of computational core were proven to work in our protocells following each step from recognition through killing. The essential point of such a strategy was that DNA RN can be used to define the logical analyzing ability and biofunction of an artificial cell. From the perspective of broader biological applications, any DNA RNs, or computational core, as herein demonstrated could theoretically interact directly with such

biological targets as nucleic acids or proteins inside the protocell. DNA RNs built into protocells can be regarded as cell-sized robots with autonomous computing capability. Therefore, this report provides a strategy for nanoengineering reliable prototypes to interact with complex chemical targets *in situ*, or even interface with existing biological RNs *in vivo*. Specifically, since DNA computational core-driven protocells can be much more deliverable than diffusible DNA RNs due to lack of spatial isolation, electrolyte-rich circulatory system including hematological system and lymphatic system seems to be a potential working environment in the future because both surface and cavity of a protocell are available for recognition and computing. Nevertheless, some problems such as proposing more active recognition mechanisms and more stable membrane structures should be subsequently solved.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

## REFERENCES

(1). Elani Y; Law RV; Ces O Phys. Chem. Chem. Phys 2015, 17 (24), 15534–15537. [PubMed: 25932977]

(2). Sunami T; Matsuura T; Suzuki H; Yomo T Methods Mol. Biol 2010, 607, 243–256. [PubMed: 20204862]

(3). Brea RJ; Hardy MD; Devaraj NK Chem. - Eur. J 2015, 21 (36), 12564–12570. [PubMed: 26149747]

(4). Elani Y; Law RV; Ces O Nat. Commun 2014, 5, 5305. [PubMed: 25351716]

(5). Kurihara K; Okura Y; Matsuo M; Toyota T; Suzuki K; Sugawara T Nat. Commun 2015, 6, 8352. [PubMed: 26418735]

(6). Kurihara K; Tamura M; Shohda K.-i.; Toyota T; Suzuki K; Sugawara T Nat. Chem 2011, 3 (10), 775–781. [PubMed: 21941249]

(7). Qiao Y; Li M; Booth R; Mann S Nat. Chem 2017, 9 (2), 110–119. [PubMed: 28282044]

(8). Takakura K; Toyota T; Sugawara T J. Am. Chem. Soc 2003, 125 (27), 8134–8140. [PubMed: 12837083]

(9). Kholodenko BN Nat. Rev. Mol. Cell Biol 2006, 7, 165–176. [PubMed: 16482094]

(10). Purvis JE; Lahav G Cell 2013, 152 (5), 945–956. [PubMed: 23452846]

(11). Elani Y; Gee A; Law RV; Ces O Chem. Sci 2013, 4 (8), 3332–3338.

(12). Peters RJ; Marguet M; Marais S; Fraaije MW; van Hest J; Lecommandoux S Angew. Chem., Int. Ed 2014, 53 (1), 146–150.

(13). Adamala KP; Martin-Alarcon DA; Guthrie-Honea KR; Boyden ES Nat. Chem 2016, 9 (5), 431–439. [PubMed: 28430194]

(14). Kamiya K; Kawano R; Osaki T; Akiyoshi K; Takeuchi S Nat. Chem 2016, 8, 881–889. [PubMed: 27554415]

(15). Burns JR; Seifert A; Fertig N; Howorka S Nat. Nanotechnol 2016, 11, 152–156. [PubMed: 26751170]

(16). Peng R; Wang H; Lyu Y; Xu L; Liu H; Kuai H; Liu Q; Tan W J. Am. Chem. Soc 2017, 139 (36), 12410–12413. [PubMed: 28841373]

(17). Lv Y; Hu R; Zhu G; Zhang X; Mei L; Liu Q; Qiu L; Wu C; Tan W Nat. Protoc 2015, 10 (10), 1508–1524. [PubMed: 26357007]

(18). Lv Y; Cui L; Peng R; Zhao Z; Qiu L; Chen H; Jin C; Zhang X-B; Tan W Anal. Chem 2015, 87 (23), 11714–11720. [PubMed: 26505212]

(19). Qian L; Winfree E Science 2011, 332 (6034), 1196–1201. [PubMed: 21636773]

(20). Seelig G; Soloveichik D; Zhang DY; Winfree E Science 2006, 314 (5805), 1585–1588. [PubMed: 17158324]

(21). Han D; Zhu Z; Wu C; Peng L; Zhou L; Gulbakan B; Zhu G; Williams KR; Tan W J. Am. Chem. Soc 2012, 134 (51), 20797–20804. [PubMed: 23194304]

(22). You M; Zhu G; Chen T; Donovan MJ; Tan W J. Am. Chem. Soc 2015, 137 (2), 667–674. [PubMed: 25361164]

(23). You M; Peng L; Shao N; Zhang L; Qiu L; Cui C; Tan W J. Am. Chem. Soc 2014, 136 (4), 1256–1259. [PubMed: 24367989]

(24). Wu Z; Liu G-Q; Yang X-L; Jiang J-H J. Am. Chem. Soc 2015, 137 (21), 6829–6836. [PubMed: 25969953]

(25). Thubagere AJ; Li W; Johnson RF; Chen Z; Doroudi S; Lee YL; Izatt G; Wittman S; Srinivas N; Woods D; et al. Science 2017, 357 (6356), eaan6558.

(26). Groves B; Chen Y-J; Zurla C; Pochekailov S; Kirschman JL; Santangelo PJ; Seelig G Nat. Nanotechnol 2016, 11, 287–294. [PubMed: 26689378]

(27). Han D; Wu C; You M; Zhang T; Wan S; Chen T; Qiu L; Zheng Z; Liang H; Tan W Nat. Chem 2015, 7 (10), 835–841. [PubMed: 26391084]

(28). Li W; Moore MJ; Vasilieva N; Sui J; Wong SK; Berne MA; Somasundaran M; Sullivan JL; Luzuriaga K; Greenough TC; et al. Nature 2003, 426 (6965), 450–454. [PubMed: 14647384]

(29). Matsuyama S; Ujike M; Morikawa S; Tashiro M; Taguchi F Proc. Natl. Acad. Sci. U. S. A 2005, 102 (35), 12543–12547. [PubMed: 16116101]

(30). Shirato K; Kawase M; Matsuyama S J. Virol 2013, 87 (23), 12552–12561. [PubMed: 24027332]

(31). Lee S; Koo H; Na JH; Lee KE; Jeong SY; Choi K; Kim SH; Kwon IC; Kim K ACS Nano 2014, 8 (5), 4257–4267. [PubMed: 24754537]

(32). Shohda K.-i.; Tamura M; Kageyama Y; Suzuki K; Suyama A; Sugawara T Soft Matter 2011, 7 (8), 3750–3753.

(33). Lakin MR; Youssef S; Polo F; Emmott S; Phillips A Bioinformatics 2011, 27 (22), 3211–3213. [PubMed: 21984756]

(34). Puglisi JD; Tinoco I Methods Enzymol 1989, 180, 304–325. [PubMed: 2482421]

(35). Yoshina-Ishii C; Boxer SG J. Am. Chem. Soc 2003, 125 (13), 3696–3697. [PubMed: 12656589]

(36). Liu H; Moynihan KD; Zheng Y; Szeto GL; Li AV; Huang B; Van Egeren DS; Park C; Irvine DJ Nature 2014, 507 (7493), 519–522. [PubMed: 24531764]

(37). Tan X; Lu X; Jia F; Liu X; Sun Y; Logan JK; Zhang K J. Am. Chem. Soc 2016, 138 (34), 10834–10837. [PubMed: 27522867]

(38). Walls AC; Tortorici MA; Bosch B-J; Frenz B; Rottier PJM; DiMaio F; Rey FA; Veesler D Nature 2016, 531 (7592), 114–117. [PubMed: 26855426]

(39). Ungar D; Hughson FM Annu. Rev. Cell Dev. Biol 2003, 19 (1), 493–517. [PubMed: 14570579]

(40). Hu C-MJ; Fang RH; Wang K-C; Luk BT; Thamphiwatana S; Dehaini D; Nguyen P; Angsantikul P; Wen CH; Kroll AV; et al. Nature 2015, 526 (7571), 118–121. [PubMed: 26374997]

(41). Chan V; Novakowski SK; Law S; Klein-Bosgoed C; Kastrup CJ Angew. Chem 2015, 127 (46), 13794–13797.

(42). Ho J-A; Wauchope R Anal. Chem 2002, 74 (7), 1493–1496. [PubMed: 12033235]

(43). Ou L-J; Liu S-J; Chu X; Shen G-L; Yu R-Q Anal. Chem 2009, 81 (23), 9664–9673. [PubMed: 19877619]

(44). Caschera F; Stano P; Luisi PL J. Colloid Interface Sci 2010, 345 (2), 561–565. [PubMed: 20181356]

(45). Chan Y-HM; van Lengerich B; Boxer SG Biointerphases 2008, 3 (2), FA17–FA21. [PubMed: 20408664]

(46). Nomura F; Inaba T; Ishikawa S; Nagata M; Takahashi S; Hotani H; Takiguchi K Proc. Natl. Acad. Sci. U. S. A 2004, 101 (10), 3420–3425. [PubMed: 14988507]

(47). Paleos CM; Tsiourvas D; Sideratou Z ChemBioChem 2011, 12 (4), 510–521. [PubMed: 21337482]

(48). Pantazatos D; MacDonald R J. Membr. Biol 1999, 170 (1), 27–38. [PubMed: 10398758]

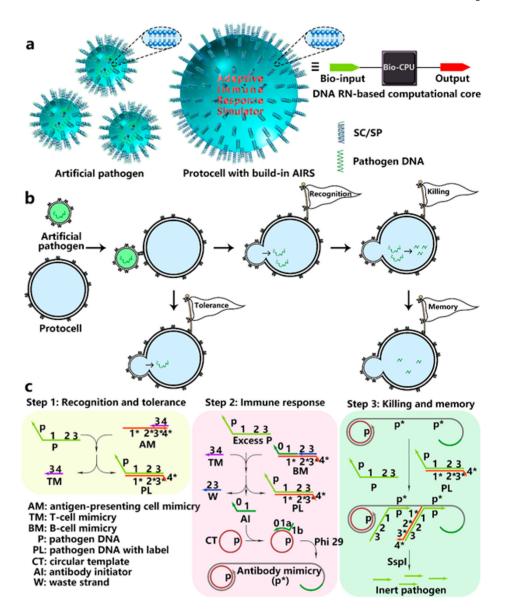(49). Stengel G; Zahn R; Höök F J. Am. Chem. Soc 2007, 129 (31), 9584–9585. [PubMed: 17629277]

**Figure 1.**
Schematic figure of protocells with built-in DNA RN-based computational core. (a) The artificial pathogen consists of a bilayer-membrane structure, loaded pathogen DNA and cholesterol-modified dsDNA inserted on the membrane. The protocell consists of a bilayer-membrane structure, encapsulated AIRS, and cholesterol-modified dsDNA on the membrane. (b) Cartoon figures that illustrate the infection mimicry between artificial pathogen and protocell, followed by triggered AIRS inside of a protocell to eliminate the infected pathogen DNA as the result of DNA RN-based computation. (c) Working principle of the DNA computational core (AIRS) built inside the protocell. Step 1: recognition and tolerance. Step 2: immune response. Step 3: killing and memory. Abbreviations used in this paper are listed in Table S2.
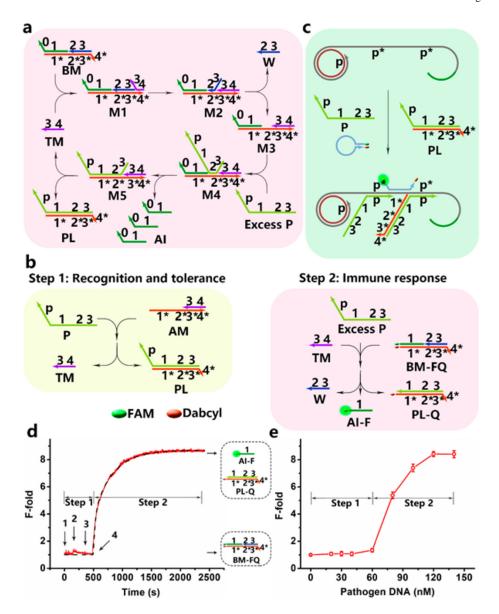
**Figure 2.**
Fluorescence signal was used to monitor the operating status of AIRS. (a) TM-catalyzed DNA strand displacement reaction. (b) Fluorescence scheme of steps 1 and 2 of AIRS. In order to give a fluorescence signal when steps 1 and 2 are triggered, an FITC-modified strand AI-F (green) was used instead of strand AI. Dabcyl was modified on the 5′-end of strand L (orange), as shown in the BM-FQ complex. Once excessive pathogen DNA triggers step 2, strand AI-F is released from the BM-FQ complex, giving a recovery signal. (c) Design of Antibody-MB to probe antibody mimicry. Loop domain of Antibody-MB is designed to be complementary with generated antibody mimicry. Consequently, Antibody-MB is opened by generated antibody mimicry through RCA to give a fluorescence signal. (d) Fluorescence kinetics with addition of different amounts of pathogen DNA. Time point 1: No pathogen DNA. Time point 2: addition of 20 nM pathogen DNA. Only step 1 was triggered. Time point 3: addition of 20 nM pathogen DNA; 40 nM of pathogen DNA in total.

Only step 1 was triggered. Time point 4: addition of 140 nM of pathogen DNA; 180 nM of pathogen DNA in total. Increasing fluorescence intensity indicates that both step 1 and step 2 were triggered. Dashed line: Visual DSD simulation result. (e) Plot of DNA-dependent fluorescence response vs concentration of pathogen. Error bars show the standard deviations of measurements taken from three independent experiments.
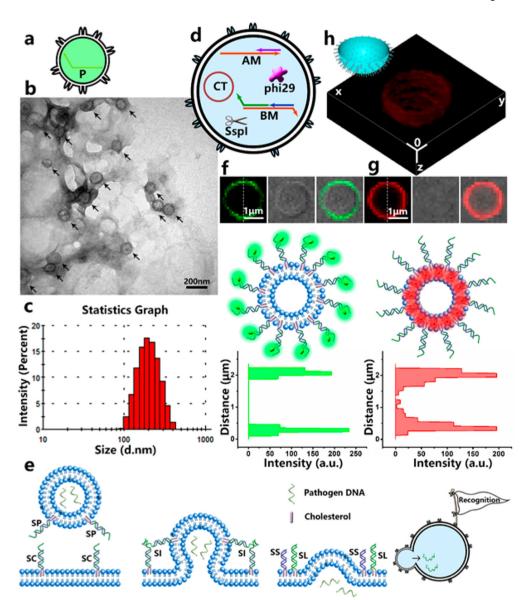
**Figure 3.**
Characterization of protocell and artificial pathogen. (a) Cartoon figure of an artificial pathogen. (b) TEM image of artificial pathogens. Pathogens are negatively stained with 2% uranyl acetate. Scale bar: 200 nm. (c) DLS determined size distribution of artificial pathogens. (d) Cartoon figure of a protocell. (e) Schematic figure of DNA strand migration as SNARE protein mimicry. (f) Fluorescence imaging of protocell. FITC-modified dsDNA SC-F was used to stain the membrane (from top left to top right: FITC channel, bright field and overlay. Scale bar: 1 $\mu$m. Middle: structural profile of a protocell stained by SC-F. Bottom: fluorescent intensity along the dashed line in FITC channel). (g) Fluorescence imaging of protocell. Nile red was used to stain the membrane (from top left to top right: Nile red channel, bright-field, and overlay. Scale bar: 1 $\mu$m. Middle: structural profile of a protocell stained by Nile red. Bottom: fluorescent intensity along the dashed line in TAMRA

channel). (h) 3D reconstructed fluorescence imaging of half a single protocell. TAMRA-modified dsDNA SC-T was used to stain the membrane. Scale bar: 500 nm.
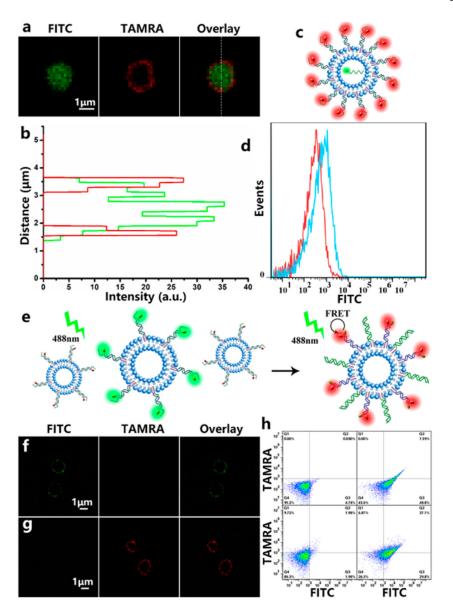
**Figure 4.**
Encapsulation of DNA strands in protocells and artificial infection mimicry. (a–c) FITC-labeled ssDNA EP-F was loaded into protocells. TAMRA-modified dsDNA SC-T was used to stain the membrane of protocell. (a) Fluorescence image of a single protocell. Scale bar: 1 $\mu$m. (b) Fluorescent intensities of red and green channels along the dashed line in panel a. Intensities in green and red channels are indicated in corresponding colors. (c) Structural profile of a protocell imaged in panel a. (d) Flow cytometry results of encapsulating strand EP-F in protocells. Blue line: protocells encapsulated with FITC-labeled ssDNA EP-F. Red line: no encapsulation. (e–g) Membrane fusion-induced FRET. (e) FITC-labeled dsDNA SC-F and TAMRA-labeled dsDNA SP-T were individually inserted onto the surfaces of protocells and artificial pathogens. FRET occurred between FITC and TAMRA as a result of the strand migration between SC-F and SP-T after artificial infection mimicry. (f) Fluorescence image of protocells before fusing with artificial pathogens. Scale bar: 1 $\mu$m. (g)

Fluorescence image of protocells fused with artificial pathogens. Scale bar: 1 $\mu$m. Only a 488 nm laser was used in this experiment, and both FITC and TAMRA signals were recorded. (h) Flow cytometry results of membrane fusion between protocells and artificial pathogens. Left top: No modification on the protocell membrane. Right top: SC-F was inserted into the membrane of protocells and gave a shift in FITC channel. Left bottom: TAMRA-labeled dsDNA SC-T was inserted into the membrane of protocells. Right bottom: SC-F-modified protocells were fused and infected by SP-T-modified artificial pathogens. Only a 488 nm laser was used, and both FITC and TAMRA signals were recorded.
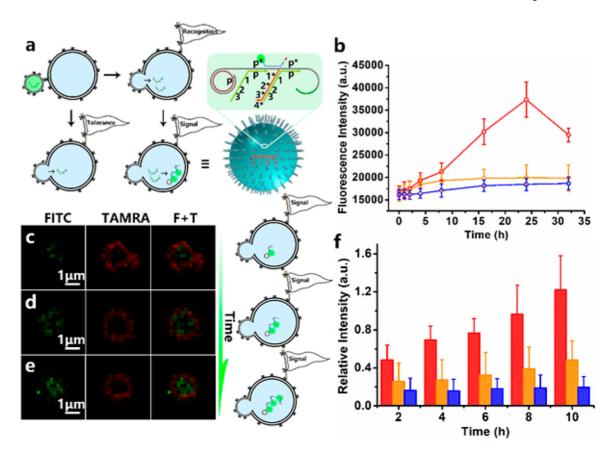
**Figure 5.**
Artificial immune response triggered by pathogen infection mimicry. (a) Schematic figure of triggered artificial immune response inside a protocell with a fluorescence signal. (b) Fluorescence kinetics of protocells after infection by different numbers of artificial pathogens. Red line: excessive pathogens; orange line: small amount of pathogen below the threshold defined in step 1; blue line: no pathogen. (c–e) Confocal imaging of different time points of triggered AIRS inside a protocell, as determined in real time: (c) 2, (d) 6, and (e) 10 h. Scale bar: 1 $\mu$m. (f) Fluorescent intensities of protocells with different responding times. Red: protocells activated by excessive artificial pathogens. Orange: protocells activated by pathogens beyond the threshold. Blue: protocells with no artificial pathogen challenge.
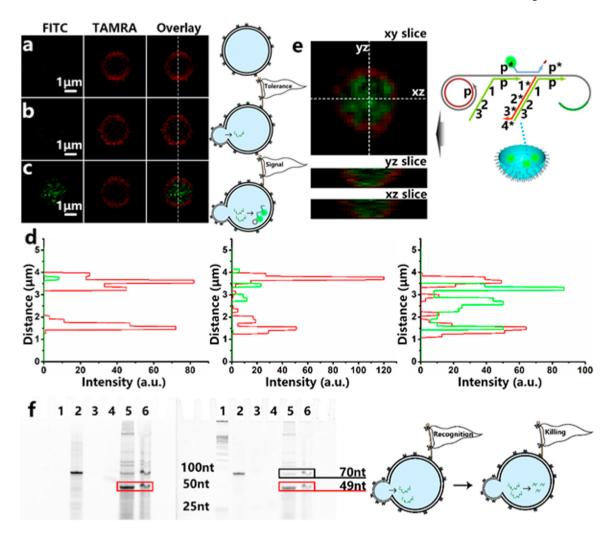
**Figure 6.**
Confocal imaging and electrophoresis results of protocells with built-in AIRS as computational core in the presence of artificial pathogenic challenge. (a) Protocell only. Encapsulated AIRS was not triggered. (b) Protocell infected with a small number of artificial pathogens. Only step 1 was triggered so that encapsulated antibody-MB was not opened. (c) Protocell infected with a large number of artificial pathogens. Encapsulated AIRS was triggered. (d) From left to right: Fluorescent intensities along the dashed lines in panels a–c. Intensities in green and red channels are indicated in corresponding colors. (e) z axis scanning and reconstruction of panel c. Lines on the xy slice indicate the positions of yz slice and xz slice. (f) SDS-PAGE image of AIRS in buffer solution and in protocell. FITC-modified pathogen DNA P–F was used in this gel to give a fluorescence signal. After recording the fluorescence signal of FITC, gel was then stained with SYBR Gold Nucleic Acid Gel Stain for one-half hour and imaged again. Left: FITC channel. Right: SYBR Gold channel. Bands in red boxes are inert pathogen DNA fragments after SspI digestion.