ACCOUNTS
of chemical research

RTRACS: A Modularized RNA-Dependent RNA Transcription System with High Programmability

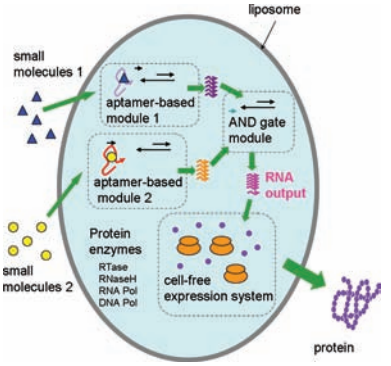
SHOTARO AYUKAWA,[†] MASAHIRO TAKINOUE,[†] AND DAISUKE KIGA*[,†,‡]

[†]*Department of Computational Intelligence and Systems Science, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8503, Japan, and* [‡]*PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*

RECEIVED ON MAY 4, 2011

CONSPECTUS

Creating artificial biological systems is an important research endeavor. Each success contributes to synthetic biology and adds to our understanding of the functioning of the biomachinery of life. In the construction of large, complex systems, a modular approach simplifies the design process: a multilayered system can be prepared by integrating simple modules. With the concept of modularity, a variety of synthetic biological systems have been constructed, both in vivo and in vitro. But to properly develop systems with desired functions that integrate multiple modules, researchers need accurate mathematical models. In this Account, we review the development of a modularized artificial biological system known as RTRACS (reverse transcription and transcription-based autonomous computing system). In addition to modularity, model-guided predictability is an important feature of RTRACS.



RTRACS has been developed as an in vitro artificial biological system through the assembly of RNA, DNA, and enzymes. A fundamental module of RTRACS receives an input RNA with a specific sequence and returns an output RNA with another specific sequence programmed in the main body, which is composed of DNA and enzymes. The conversion of the input RNA to the output RNA is achieved through a series of programmed reactions performed by the components assembled in the module. Through the substitution of a subset of components, a module that performs the AND operation was constructed. Other logical operations could be constructed with RTRACS modules. An integration of RTRACS modules has allowed the theoretical design of more complex functions, such as oscillation. The operations of these RTRACS modules were readily predicted with a numerical simulation based on a mathematical model using realistic parameters.

RTRACS has the potential to model highly complex systems that function like a living cell. RTRACS was designed to be integrated with other molecules or molecular devices, for example, aptazymes, cell-free expression systems, and liposomes. For the integration of these new modules, the quantitative controls of each module based on the numerical simulation will be instructive. The capabilities of RTRACS promise to provide models of complex biomolecular systems that are able to detect the environment, assess the situation, and react to overcome the situation. Such a smart biomolecular system could be useful in many applications, such as drug delivery systems.

1. Introduction: Modularity in Synthetic Biological Systems

The constructive approach is important in synthetic biology, because it aims to understand the processing of complex biomolecular systems. An effective method for the construction of artificial biological systems is the integration of simple modules composed of biomolecules such as DNA, RNA, and proteins[1] (Figures 1A,B and 2A−C). Two modules can be integrated when the output of a module and the input of the other module are the same. A complex multilayered system can be constructed by integrating simple modules. The analysis of the modules using mathematical

Published on the Web 10/19/2011 www.pubs.acs.org/accounts
10.1021/ar200128b

modeling based on realistic parameters plays an important role in module integration.[2] In this Account, we describe the importance of modularization and modeling in the construction of biological systems within the context of the RTRACS (reverse transcription and transcription-based autonomous computing system) approach. RTRACS is an in vitro modularized RNA-dependent RNA transcription system whose modules (Figure 1C) are designed to be integrated in several ways.

In synthetic biology, a module can be defined as a biomolecular unit that receives an input and returns an output. In an artificial genetic circuit encoded in DNA, each gene is considered to be a module. For example, a repression module composed of a LacI regulated promoter, a ribosome-binding site, and a GFP ORF turns off generation of GFP as an output in response to the input of LacI repressor (Figure 1B). In synthetic biology, many artificial biological modules have been constructed in vitro as well as in vivo. An aminoacyl-tRNA synthetase (aaRS) can be considered as a module that receives a tRNA as an input and returns an aminoacyl-tRNA as an output (Figure 1A). An aaRS that is modified to attach an unnatural amino acid to tRNA was used a module to expand the genetic code in vitro.[3] Most nucleic acid computers are also composed of modules where each receives an input nucleic acid strand and returns an output strand. Soloveichik et al. designed DNA-based modules that generate an output DNA strand via a strand displacement reaction initiated by an input DNA strand.[4]

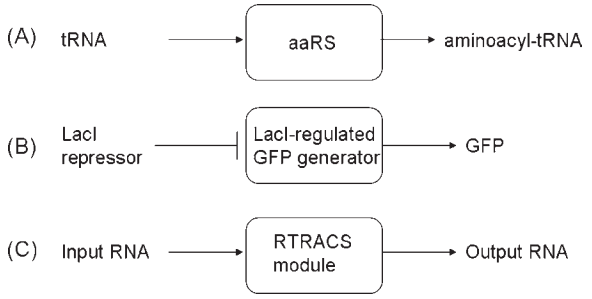


FIGURE 1. Modules in biological systems. (A) An aminoacyl-tRNA synthetase (aaRS) receives a tRNA and returns an aminoacyl-tRNA. (B) A LacI-regulated GFP generator receives a LacI repressor and turns off a GFP. (C) An RTRACS module receives an input RNA and returns an output RNA.

Many artificial biological systems have been constructed in synthetic biology by integrating simple modules (Figure 2). A toggle switch working inside *E. coli* was produced by integrating two repression modules in a form of mutual inhibition, where each was composed of a repressor protein and the corresponding regulatory expression unit (Figure 2A).[5] Similarly, the integration of three repression modules into an inhibition-loop form was employed in the construction of an oscillator system (Figure 2B).[6] A variety of nucleic acid-based systems have been constructed, including logic gates,[7,8] a toggle switch,[9] and oscillators,[10,11] by integrating nucleic acid-based modules where each is composed of designed nucleic acid strands and sometimes enzymes.

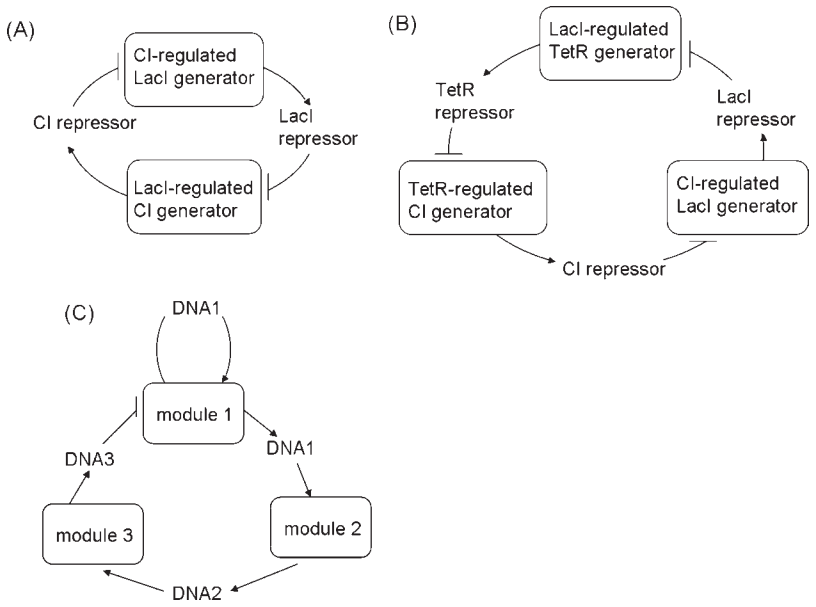


FIGURE 2. Integrations of modules in biological systems. (A) A genetic toggle switch composed of two repression modules.[5] (B) A genetic oscillator composed of three repression modules.[6] (C) An oscillator composed of nucleic acid-based modules.[11]

The integration of simple modules generates upper layer modules. A population control system[12] was constructed by introducing a cell−cell communication gene to function as a module into living cells. Each cell that contained the communication module also functioned as a module in an upper layer system that could control the population dynamics. Integration of simple modules by Seelig et al. generated a five-layer logic gate that received six different nucleic acid strands.[7]

Mathematical models can be used to analyze the dynamics of integrated modules in a large system. Such predictions are effective in excluding unrealistic integration of modules from a large range of possible combinatorial patterns. A difficult problem in the integration of modules is to harmonize module functions, such as the output production rate in response to the input. Mathematical models based on realistic parameters supported the construction of the previously mentioned toggle switches[5,9] and oscillators[6,10,11] both in vivo and in vitro. Mathematical modeling is indispensable in synthetic biology when implementing a designed system.

Nucleic acid computing, which is an attractive topic in the field of in vitro synthetic biology, also uses the constructive approaches based on modularization, module integration, multilevel modularity, and mathematical modeling. The specificity of hybridization based on Watson−Crick complementarity makes nucleic acids a suitable material to use as the main component. In addition, the catalytic activities of nucleic acid strands provide nucleic acid computers with higher order operations. Several types of nucleic acid computers have been constructed by assembling nucleic acid strands with designed sequences.[13,14] The specificity of strand complementarity enabled the construction of hybridization-based computers with a variety of functions, such as combinatorial search[15] and logic gates.[7,16] Nucleic acid computers that utilize the catalytic activities of enzymes[17,18] or deoxyribozymes[19] possess the advantage of amplification and/or conversion of the input or output signals. Mathematical models containing the parameters of the respective components were used in the construction of some of these nucleic acid computers to confirm whether the computers would work as planned, as is the case for in vivo synthetic biology. These properties of nucleic acids are also utilized to the construction of nanostructures or nanomachines.[20]

In this Account, we describe a type of modular nucleic acid computer known as RTRACS (Reverse-transcription and TRanscription-based Autonomous Computing System). Each RTRACS module receives an input RNA and returns an output RNA (Figure 1C) as an enzyme receives a substrate and returns a product (Figure 1A). In each module, the conversion of an input RNA sequence to an output RNA sequence is performed isothermally in a one-pot reaction via a series of reactions initiated by hybridization between an input RNA strand and a DNA strand containing the template region for transcription of output RNA. Other modules were constructed to perform logic gate operations by the substitution of subsets of the components of a module. In a system composed of integrated RTRACS modules, a pair of modules can be connected because both modules use RNA as the input and output. In addition, the equality of the working temperature of RTRACS modules assures the connectivity of the modules. Moreover, the dynamics of all the modules and that of the module-integrated system could be predicted by mathematical modeling, because all of the components and reactions in each module are identified. RTRACS also has the potential to be integrated with other biomolecular devices, such as aptazyme-based devices that receive specific molecular inputs. In the next section, we will discuss the characteristics and applications of RTRACS.

2. Development of RTRACS

2.1. Fundamental RTRACS Modules. RTRACS consists of individual modules which convert an input RNA strand with a specific sequence to an output RNA strand with another sequence (Figure 1C). In the design of a module,[21] an RNA strand is synthesized as an output after a logical operation that is achieved by a series of autonomously regulated hybridization and enzymatic reactions. This element contains a de novo combination of a reverse transcriptase, a DNA polymerase, an RNA polymerase, a ribonulease H, a primer DNA, and a converter DNA. The converter DNA contains the single-stranded complementary sequences of the input RNA and a T7 promoter as well as the double-stranded template sequence of the output RNA. The working temperature range for a module is dependent on the temperature range of the enzymes. Figure 3 shows the reaction scheme of the RTRACS module. First, the converter DNA hybridizes to the input RNA (Figure 3(1)). Using the hybridized converter DNA as a primer and the input RNA as a template, the reverse transcriptase synthesizes the complementary DNA of the input RNA (Figure 3(2)). The synthesized complementary DNA becomes single-stranded through the digestion of the input RNA by the ribonuclease H (Figure 3(3)). Then the primer DNA hybridizes to the specific site on the complementary DNA (Figure 3(4)). Subsequently, the complementary DNA becomes double-stranded through the extension with the DNA polymerase (Figure 3(5)). This

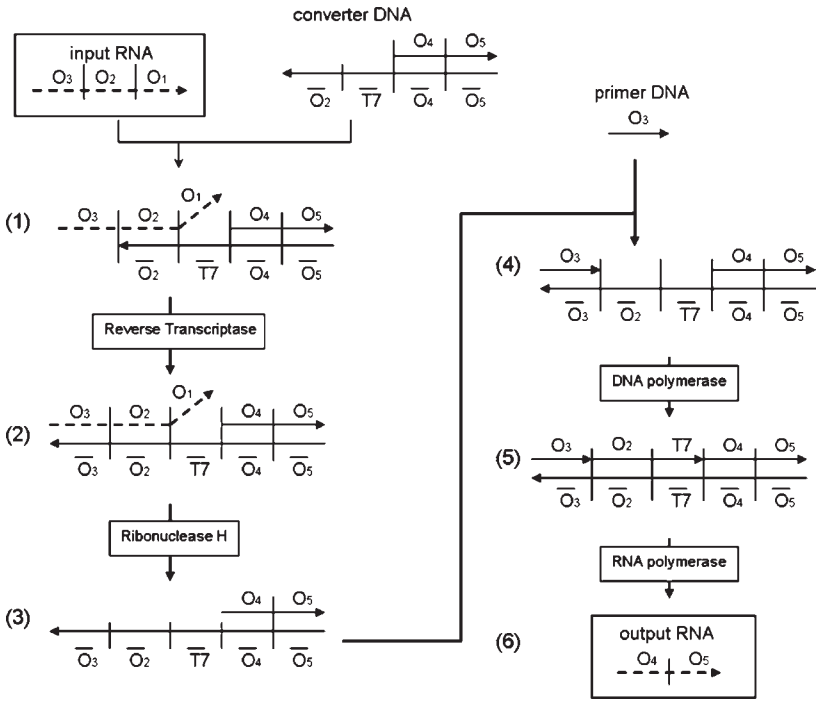


FIGURE 3. Reaction scheme of the fundamental module of RTRACS. Solid arrows indicate DNA strands. Dashed arrows indicate RNA strands. An output RNA is produced only when an input RNA exists via the following multistep isothermal reactions. (1) The converter DNA hybridizes to the input RNA. (2) The reverse transcriptase synthesizes the complementary DNA of the input RNA using the hybridized converter DNA as a primer and the input RNA as a template. (3) The synthesized complementary DNA becomes single-stranded via the digestion of the input RNA by the ribonuclease H. (4) The primer DNA hybridizes to the specific site on the complementary DNA. (5) The complementary DNA, which includes a promoter sequence, becomes double-stranded through extension with reverse transcriptase. (6) The RNA polymerase recognizes the double-stranded T7 promoter and transcribes the output RNA strand. The arrowheads indicate 3′-termini of the nucleic acid strands. The DNA and RNA strands are composed of uniform-length orthonormal sequences from O_1 to O_5 and their complementary sequences from $\bar{O}_1$ to $\bar{O}_5$.

process also makes the promoter sequence on the converter DNA double-stranded (Figure 3(5)). Finally, the RNA polymerase recognizes the double-stranded T7 promoter and transcribes the output RNA strands (Figure 3(6)). On the whole, these reactions work as an encoder that generates the output RNA only when the corresponding input RNA exits.

The programmed function of an RTRACS module was confirmed in experiments using two modules where each was designed to receive a specific input RNA. Module 1 and module 2 contain different converter DNAs each of which has an input-RNA binding sequence for RNA A and RNA B, respectively. An output RNA was successfully produced when RNA A was used as the input for module 1 (Figure 4A). Conversely, in module 2, the output RNA was produced when RNA B was used as the input (Figure 4B).

2.2. Development of Orthonormal Sequences for Construction of RTRACS. Programmability based on sequence complementarity makes nucleic acids useful materials, but problems arising from mishybridization and self-folding of nucleic acid strands need to be addressed. One approach to this problem is the development of a set of "orthonormal sequences" with very little unintentional partial complementarity. Orthonormal sequences were designed to have uniform lengths and melting points with no potential for mishybridization and stable self-folding.[22] The accuracy and efficiency of the interaction between strands programmed to form base-parings can be improved by encoding the sequences of the component strands as orthonormal sequences. Therefore, the use of orthonormal sequences reduces undesirable interactions among strands and simplifies the construction of RTRACS modules.

2.3. Logical Operations with RTRACS. By using the RTRACS mechanism, Takinoue et al. designed and constructed a module that performs the AND gate operation (Figure 5).[23] In this mechanism, the output RNA strand R_C is generated only when two different input RNA strands (R_A and R_B) coexist. Logical values of 0 and 1 are represented by the absence and presence of RNA strands, respectively. The presence of an input RNA strand R_A means a logical value of 1 for Input A, while the absence of the input RNA means a logical value of 0 for Input A. Similarly, the presence and absence of other input and output RNAs mean corresponding

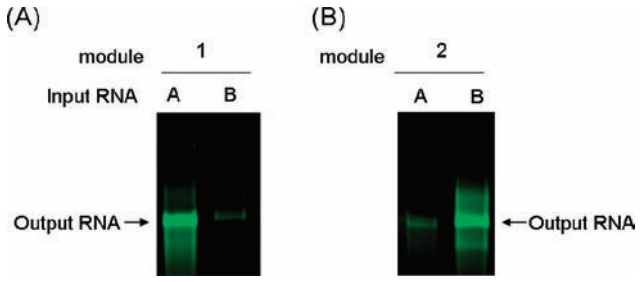


FIGURE 4. Detection of the output RNA produced in response to the input RNA in RTRACS modules. Module 1 was designed to receive input RNA A and not input RNA B, because the converter DNA only hybridizes to input RNA A and vice versa. Both modules contained *Thermus thermophilus* ribonuclease H, avian myeloblastosis virus (AMV) reverse transcriptase that functioned as both a reverse transcriptase and a DNA polymerase, and thermostable T7 RNA polymerase. Module operations were performed for 60 min at 50 °C. The reaction volume was 10 μL. The band intensity of the output RNA was measured using a fluoroimage analyzer, FLA-5100 (Fuji Film, Japan). (A) The production of the output RNA in module 1 in response to the input of RNA A and RNA B. RNA A induced the production of 19-fold greater amounts of output RNA compared with RNA B, although a leaky generation occurred with the input of RNA B. (B) The production of the output RNA in module 2 in response to the input of RNA A and RNA B. RNA B induced the production of 9.4-fold greater amounts of output RNA compared with RNA A, although a leaky generation occurred with the input of RNA A. Copyright 2011 YODOSHA CO., LTD.

logical values (Table 1). At the beginning of the AND gate reaction, R_A hybridizes with primer DNA P1 to form a DNA–RNA hybrid P2. P2 is then converted to a full DNA–RNA hybrid P3 by extension with the reverse transcriptase. Thereafter, the RNA strand of P3 is degraded by ribonuclease H and a single-stranded DNA P4 is formed. Then P4 hybridizes with the input RNA R_B to form a DNA–RNA hybrid P5. Subsequently, P5 is converted to P6 by extension with the reverse transcriptase. P6 is then converted to P7 by degradation by ribonuclease H. Thereafter, P7 hybridizes with a converter DNA Q1 to form a partial DNA–DNA hybrid Q2. Then the DNA polymerase, using Q2 as a substrate, produces a DNA–DNA hybrid Q3, in which the T7 promoter sequence becomes double-stranded. Finally, the T7 RNA polymerase recognizes the double-stranded T7 promoter and transcribes the subsequent sequence to produce the output RNA R_C. With this scheme, the output RNA strand R_C is produced only when both of the input RNA strands R_A and R_B are present.

In the above AND gate module constructed by Takinoue et al., only an input combination of logical values of 1 for both inputs was designed to produce an output RNA. However, a low rate of output RNA R_C production occurred in the other three input combinations. The regulatory range of this AND gate was evaluated based on the band intensity of

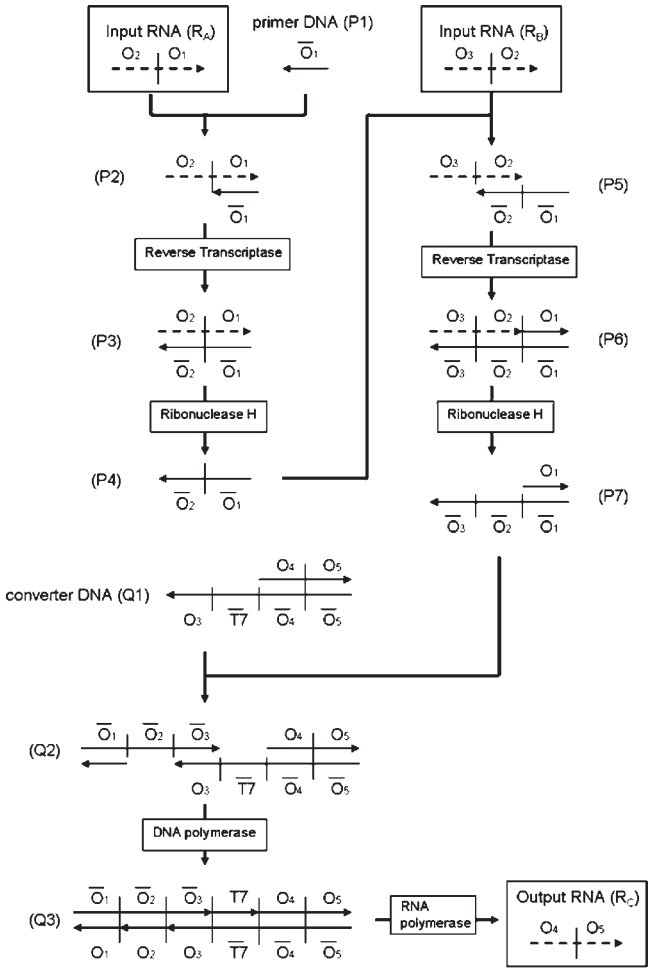


FIGURE 5. Reaction scheme of an AND gate module constructed by Takinoue et al.[23] The DNA and RNA strands are composed of orthonormal sequences from O_1 to O_5 and their complementary sequences from $\bar{O}_1$ to $\bar{O}_5$. The reaction proceeds to the end and output RNA R_C is produced only when input RNAs, R_A and R_B, exist. When only one or none of the input RNAs exists, no output RNA is produced.

TABLE 1. Evaluation of the AND Gate Constructed by Takinoue et al.[23] [a]

logical values	$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$
input combination	$\begin{bmatrix}R_A\\R_B\end{bmatrix}$	$\begin{bmatrix}R_A\\\text{no RNA}\end{bmatrix}$	$\begin{bmatrix}\text{no RNA}\\R_B\end{bmatrix}$	$\begin{bmatrix}\text{no RNA}\\\text{no RNA}\end{bmatrix}$
output C	R_C	no RNA	no RNA	no RNA
band intensity of R_C (a.u.)	1200	160	77	76
regulatory range	1	0.13	0.064	0.063

[a] The band intensities of R_C were measured from the gel shown in ref 23. The regulatory range was obtained by dividing the band intensity of R_C for each input combination by that for input combination of R_A and R_B.

R_C for each input combination. The ranges of the input combinations, R_A-no RNA, no RNA-R_B, and no RNA-no RNA were found to be 0.13-, 0.064-, and 0.063-fold that of the output RNA R_C, respectively (Table 1).

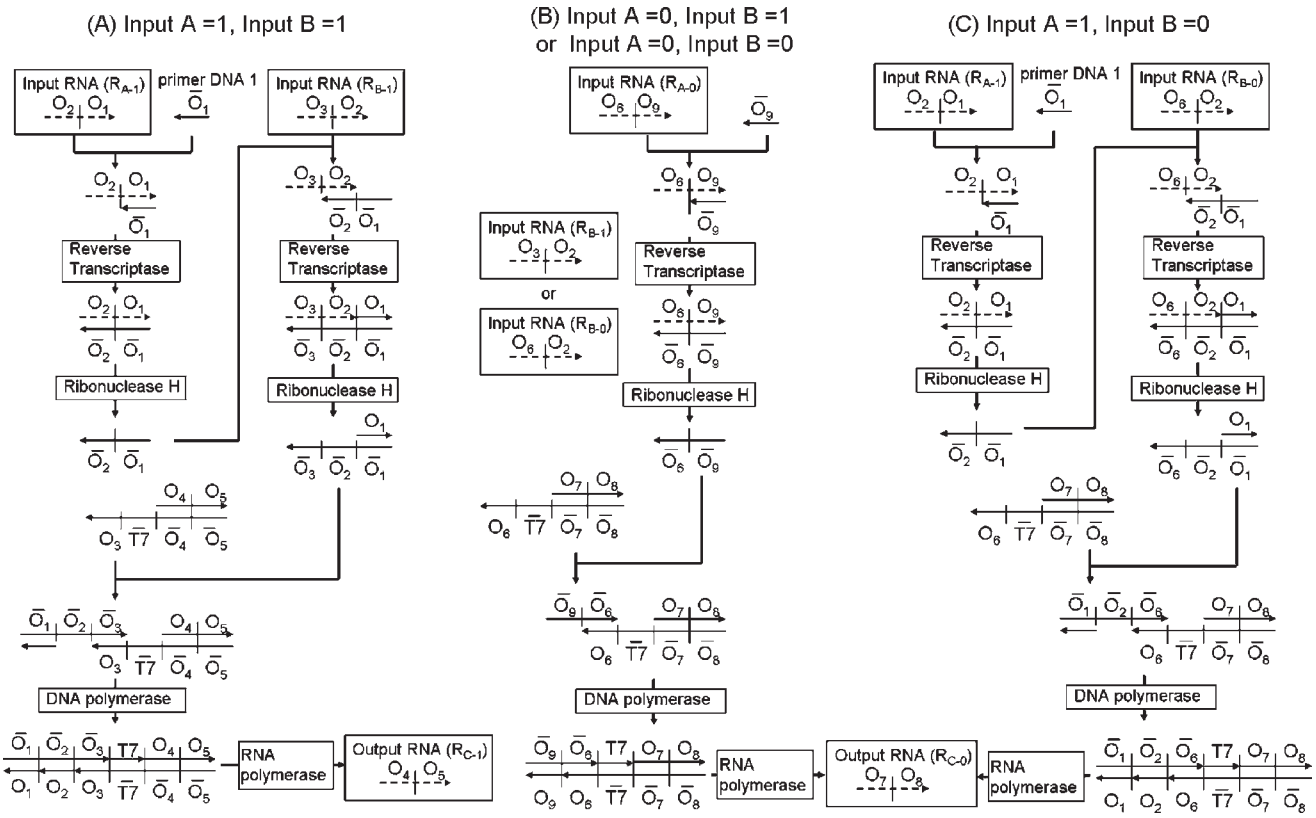


FIGURE 6. Reaction scheme of the improved AND gate module constructed by Sakai et al.[24] In this AND gate, the output RNA strands are produced with all input combinations, although the sequences of the outputs are different depending on the logical value of the output. For all input combinations, appropriate substrates of all enzymes exist. (A) The reaction occurs when both Input A and Input B are a logical value of 1. Input combination R_{A-1}-R_{B-1} induces a series of reactions to produce output RNA R_{C-1}, which has a logical value of 1 as an output. (B) The reaction occurs when Input A is a logical value of 0 and Input B is a logical value of 1 or 0. Two input combinations, including R_{A-0}, induce a different series of reactions to produce the output RNA, R_{C-0}, which has a logical value of 0 as an output. (C) The reaction occurs when Input A is a logical value of 1 and Input B is a logical value of 0. Input combination R_{A-1}-R_{B-0} induces another series of reactions to produce output RNA R_{C-0}, which has a logical value of 0 as an output. The regulatory range of the AND gate is evaluated based on the ratio of R_{C-1} to R_{C-0}. The DNA and RNA strands are composed of orthonormal sequences from O_1 to O_9 and their complementary sequences from $\bar{O}_1$ to $\bar{O}_9$.

TABLE 2. Evaluation of the AND Gate Constructed by Sakai et al.[24 a]

logical values	$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	
input combination	$\begin{bmatrix}R_{A-1}\\R_{B-1}\end{bmatrix}$	$\begin{bmatrix}R_{A-1}\\R_{B-0}\end{bmatrix}$	$\begin{bmatrix}R_{A-0}\\R_{B-1}\end{bmatrix}$	$\begin{bmatrix}R_{A-0}\\R_{B-0}\end{bmatrix}$	$\begin{bmatrix}\text{no RNA}\\\text{no RNA}\end{bmatrix}$
band intensity of R_{C-1} (a.u.)	150000	20000	12000	15000	31000
band intensity of R_{C-0} (a.u.)	6500	210000	240000	250000	9500
ratio of R_{C-1} to R_{C-0}	23	0.095	0.050	0.060	3.3
regulatory range	1	0.0041	0.0022	0.0026	

[a]The band intensities of R_{C-1} and R_{C-0} were measured from the gel shown in ref **24**. The regulatory range was obtained by dividing the ratio of R_{C-1} to R_{C-0} for each input combination by that for input combination of R_{A-1} and R_{B-1}.

In contrast, output RNA strands were produced in all four input combinations in an improved AND gate module, although the sequence of the RNA strands differed depending on the logical values of the output (Figure 6).[24] In this implementation, the presence of the input RNA strands R_{A-0} and R_{A-1} means a logical value of 0 and a logical value of 1 of Input A (Table 2). Correspondingly, the presence of the input RNA strands, R_{B-0} and R_{B-1}, means a logical value of 0 and a logical value of 1 of Input B. When input RNA strands R_{A-1} and R_{B-1} coexist, an output RNA strand R_{C-1} is produced by a series of reactions (Figure 6A). Two combinations including R_{A-0} were designed to induce a different series of reactions to

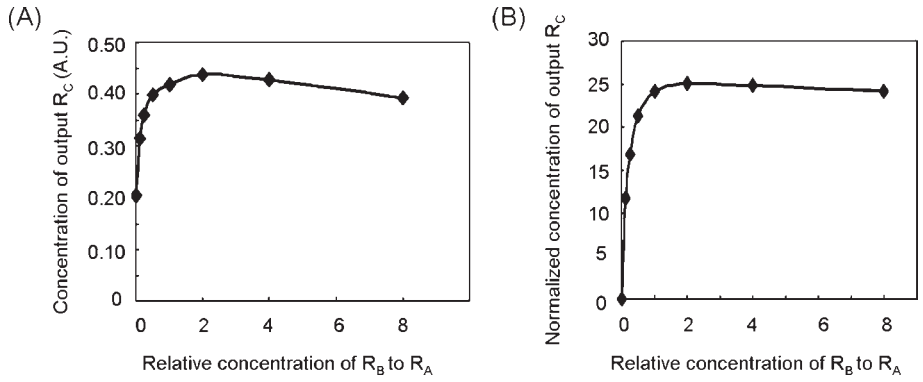


FIGURE 7. Results of the AND gate operation obtained by in vitro experiments and numerical simulations.[23] The AND gate RTRACS module shown in Figure 5 was used, and the dependence of the output RNA R_C generation on the relative concentration of the input RNA R_B to R_A was measured using an in vitro experiment (A) and a numerical simulation (B). In panel (A), the amount of the output RNA R_C was detected by using a molecular beacon that specifically hybridizes to R_C. About 0.18 of the background signal was generated by the molecular beacon without any enzymes and input RNA. The output RNA R_C generated in response to varying concentrations of the input RNA R_B at a fixed concentration of the other input, RNA R_A, is shown. The *x*-axis represents the concentration of input RNA R_B, while the *y*-axis represents the concentration of the output RNA R_C.

produce the other output RNA, namely, $R_{C\text{-}0}$ (Figure 6B). Similarly, the combination of $R_{A\text{-}1}$ and $R_{B\text{-}0}$ was designed to induce another series of reactions to produce $R_{C\text{-}0}$ (Figure 6C). The regulatory range of this AND gate was evaluated based on the ratio of $R_{C\text{-}1}$ to $R_{C\text{-}0}$. As shown in Table 2, the range of input combination $R_{A\text{-}1}$-$R_{B\text{-}1}$ was much higher than those of the other three input combinations. Thus, the regulatory range of this AND gate was improved compared with the previous gate, mainly by the association of an RNA strand with an output logical value of 0. Association of RNA strands with an input logical value of 0 also partially contributed to the improvement of the regulatory range. The band intensity of $R_{C\text{-}1}$ in the input combination of $R_{A\text{-}0}$- $R_{B\text{-}0}$ was 2-fold lower than that in the absence of both input RNAs. A decrease in excess enzymes was thought to have caused the reduction, because excess enzymes could interact with unexpected assemblies produced by mishybridization of strands in the absence of the correct substrate, although such assemblies would have lower affinities for the enzymes than the correct substrates.

The association of RNA strands with a logical value of 0 was also proposed in the design of logic gates other than the previously constructed AND and YES gate modules. The association of RNA strands with a logical value of 0 allows the construction of a NOT gate. One of a pair of YES gates produces an output RNA strand whose logical value is the opposite to that for the input RNA strand. The improved AND gate can be viewed as a NAND gate when $R_{C\text{-}0}$ and $R_{C\text{-}1}$ represent a logical value of 1 and a logical value of 0 for output C, respectively.

2.4. Mathematical Modeling of RTRACS. Because the AND gate module of RTRACS is composed of a series of DNA−DNA hybridizations, DNA−RNA hybridizations, and multiple enzymatic reactions as described above, this logic gate is not a digital gate, but an analog gate. Thus, an input−output response of this RTRACS logic gate can be described using a set of differential equations.

Here, for simplicity, hybridizations are assumed to be described by the two-state model as follows:

$$X + \bar{X} \rightleftarrows X\bar{X}$$

$$\frac{d[X\bar{X}]}{dt} = k_{on}[X][\bar{X}] - k_{off}[X\bar{X}]$$

where $[X]$ and $[\bar{X}]$ indicate concentrations of a DNA/RNA strand and its complementary strand, and k_{on} and k_{off} are forward and backward rate constants, respectively. In addition, enzymatic reactions are assumed to conform to the Michaelis−Menten model as follows:

$$S + E \rightleftarrows SE \rightarrow P + E$$

$$\frac{d[P]}{dt} = \frac{k_{cat}[E]_T[S]}{K_m + [S]}$$

where [P] and [S] represent the concentrations of a reaction product and a reactive substrate, respectively, $[E]_T$ represents the total concentration of an enzyme, and k_{cat} and K_m indicate the enzymatic reaction constant and the Michaelis−Menten constant of the enzyme, respectively. Under these assumptions, multiple differential equations corresponding to each reaction step are generated. Figure 7A shows the output RNA R_C generation in response to the varying concentrations of the input RNA R_B at a fixed

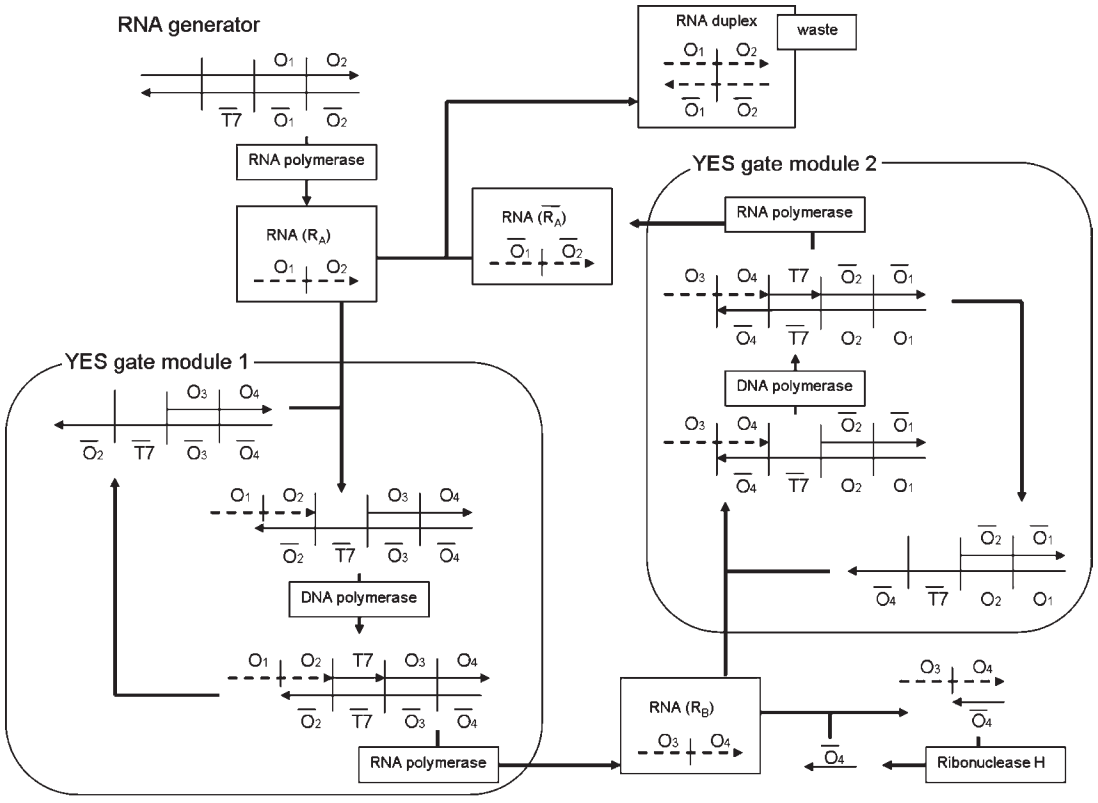


FIGURE 8. RNA oscillator composed of integrated RTRACS modules.[25] The DNA and RNA strands are composed of the orthonormal sequences from O_1 to O_4, and their complementary sequences from $\bar{O}_1$ to $\bar{O}_4$. An RNA oscillator was designed by the integration of two YES gate RTRACS modules and an RNA generator. Each YES gate module generates a specific output RNA in response to a specific input RNA. An RNA generator constitutively generates a specific output RNA..

concentration of another input, RNA R_A, in experiments; Figure 7B shows that in numerical simulations. An extreme excess of R_B over R_A was expected to reduce the R_C production rate because the binding between the intermediate product P7 and the excess R_B would inhibit the binding of P7 to the converter DNA Q1. In both results, a 2-fold excess of R_B decreased the R_C production rate. These results demonstrate that the numerical simulation result of RTRACS reproduces the result of the in vitro experiments of RTRACS qualitatively and relatively quantitatively. It is notable that the dependence of the output generation on the input concentrations is well described by the mathematical model. Thus, this result demonstrates that the input−output (I/O) response of a module of RTRACS can be well predicted and the mathematical model can be used for the design of a new RTRACS reaction circuit.

2.5. Design and Simulation of an Oscillator by Integration of RTRACS Modules. RTRACS was designed to be built up to perform desired functions by integration of modules. With this concept, Takinoue et al. proposed an RNA oscillator in which an oscillation is generated by a feedback reaction comprising two activation reactions and one inhibition reaction (Figure 8).[25] The feedback reaction is performed with three specific RNA strands, R_A, R_B, and $\bar{R}_A$. R_A is generated from an RNA generator. R_B is generated by the sequence conversion reaction of YES gate module 1 initiated by the input of R_A. $\bar{R}_A$ is likewise generated by the sequence conversion reaction of YES gate module 2 initiated by the input of R_B. Since the sequence of $\bar{R}_A$ is complementary to the sequence of R_A, the generation of $\bar{R}_A$ inactivates YES gate module 1 through hybridization between R_A and $\bar{R}_A$. Even though the RNA oscillator has not been constructed yet, its simulation results will provide important information for the construction.

2.6. RTRACS Functions As a Biomolecular Computer. From an informational viewpoint, RTRACS is considered as an autonomous biomolecular computer capable of converting an RNA sequence into another RNA sequence as a result of computation. The history of biomolecular computers began with the development of a DNA computer by Adleman to solve a mathematical problem.[15] Even though a

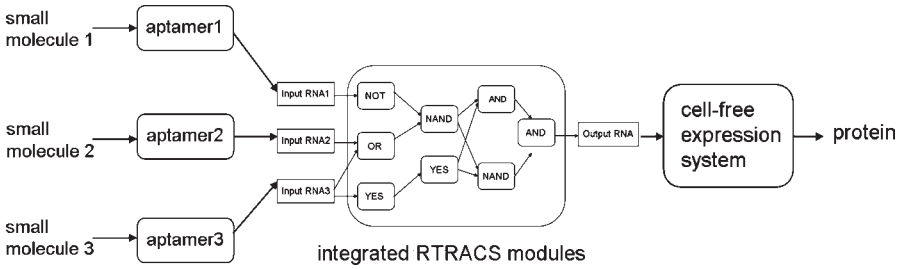


FIGURE 9. Expansion of RTRACS by integration with other biomolecular devices. Aptamer-based devices and a cell-free expression system are integrated with RTRACS to construct a system that recognizes the environment through molecular detection and produces appropriate proteins.

biomolecular computer was originally envisioned as a massive parallel computer for solving challenging mathematical problems, it has proven difficult to overcome the convenience of using silicon-based computers. In addition to the massive parallelism, biomolecular computers have been considered to have the ability to interface directly with biological systems.[26] As other biomolecular computers do, RTRACS has substantial advantages in biological applications including genetic diagnosis.[24]

3. Expansion of RTRACS

In this section, we discuss the expansion of the input and output of RTRACS. For example, RTRACS modules have the potential to be integrated with an aptamer or aptazyme-based device that generates the specific RNA strand in response to the input of the target molecule. Considering the variety of RNA aptamers,[27] RNA aptamer-based devices have wider applications than a DNA-aptamer based device.[28] For conversion of an output RNA into a protein, a cell-free expression system[29] is appropriate for integration with RTRACS. Since the sequence of the output RNA is exchangeable, an mRNA of a specific protein could be used as the output RNA. By using a cell-free expression system, the mRNA generated by an RTRACS operation could be translated into the output protein. Other than the expansion of the input and output, a liposome that works as a cell-sized container could be integrated with RTRACS. The reaction conditions of RTRACS modules and the devices must be adjusted during device integrations. For example, an aptamer-based device[28] and a cell-free translation system[30] had similar magnesium concentration requirements. When adjusting the temperature condition of the translation system, RTRACS would be required to work at 37 °C. Such temperature adjustments would be easy, because all the enzyme activities required for RTRACS could be provided by various enzymes including the same ones used with RTRACS at 50 °C, as described in this Account. In an RTRACS module, the RNA could provide an I/O interface for other devices and other RTRACS modules, as mentioned in the Introduction (Figure 9). These device integrations have the potential to produce cell-like systems that recognize their environment through molecular detection so that appropriate proteins could be produced.

4. Advantages and a Disadvantage of RTRACS

One advantage of RTRACS over other in vitro synthetic systems is that the operation of RTRACS module can be changed easily by reassembling the nucleic-acid components, which is a common strategy in the construction of nucleic acid computers. To construct AND gate modules, the nucleic-acid components of a fundamental RTRACS module that receives one input RNA strand were altered without changing the protein enzymes and reaction conditions. Compared to enzymes and reaction conditions, it is easier to change nucleic-acid components because their functionality can be controlled by redesign of their sequences.

The use of RNA is a novel aspect of RTRACS compared with other nucleic acid computers. RTRACS uses RNA strands as inputs and outputs, while the main body of an RTRACS module is composed of DNA and enzymes. In contrast, most nucleic acid computers use DNA strands as inputs and/or outputs. In a large system containing a cascade of various steps of reactions, the later reactions in a cascade may be confused by the habitual presence of input molecules used in earlier reactions. In reality, living cells use DNA as a permanent form of genetic information, while RNA is used as a transient signal for protein expression and its control. We think that this design concept of using RNA as a transient signal is unique to RTRACS. In addition, the use of RNA as an output allows RTRACS to be integrated with cell-free translation or generation of RNA aptamers, as mentioned above.

RTRACS also shares a potential disadvantage with other enzyme-based nucleic acid computers containing protein

enzymes with key roles in amplification and conversion of signals. Protein enzymes provide a broad range of catalytic activities for biomolecular computers, including natural organisms that perform complex operations, but the instability of protein enzymes results in a short life expectancy for synthetic cell-free systems. Given that living organisms use nucleic acid and proteins in their fundamental systems from the central dogma to cell−cell communication, there must be solutions for solving this problem in RTRACS.

5. Conclusions and Perspectives

In this Account, we described the importance of modularity in biological systems by focusing on the development of RTRACS. RTRACS is composed of modules, each of which receives an input RNA and returns an output RNA. A module of RTRACS was altered to perform the AND logic operation. It is also feasible to realize other logical operations with the RTRACS modules. Moreover, RTRACS modules could be integrated with each other to perform much more complex operations as desired. The development of designed functions of the integrated system is enabled by quantitative control based on mathematical modeling. In addition, the integration of RTRACS modules with other biomolecular modules such as aptamer-based devices, cell-free protein expression systems, and liposomes is being pursued. These characteristics of RTRACS are expected to lead to the development of a complex biomolecular system that can detect changes in the time pattern of lesion markers inside a tissue and to release the optimal amount of medicine in an appropriate time pattern. Such a complex biomolecular system could be applied to next-generation drug delivery systems.

We thank the other contributors, especially Professor Akira Suyama, Dr. Koh-ichiroh Shohda, Ms. Yoko Sakai, and Mr. Young-Hun Lee, for development of RTRACS. We also gratefully acknowledge the support by the SENTAN program of JST for D.K. and by the KAKENHI programs from JSPS and MEXT (09J09649 for S.A., 22220001 for M.T., 14085203, 17700291, 20650044, 21650018, and 23119005 for D.K.).

BIOGRAPHICAL INFORMATION

Shotaro Ayukawa was born in 1984 in Japan. He received his B. S. in Bioscience and M.S. in Arts from Tokyo Institute of Technology. He is currently a graduate student in the Kiga laboratory in the Department of computational intelligence and systems science at the Tokyo Institute of Technology.

Masahiro Takinoue was born in 1978 in Japan. He is a physicist studying life systems on the basis of biophysics, nonlinear nonequilibrium physics, and microfluidics. He received a B.S. in Physics in 2002 and a Ph.D. in Physics in 2007 from The University of Tokyo. He served as a postdoctoral fellow at the Kyoto University, and an assistant professor at the University of Tokyo and has been an associate professor of Tokyo Institute of Technology since 2011. His current research interests are artificial cell models, nonequilibrium active matter, DNA nanoengineering, and molecular robotics.

Daisuke Kiga was born in 1971 in Japan. His research interests include artificial genetic circuits, engineering of genetic codes, and in vitro synthetic biology. He received a B.S. in Biophysics and Biochemistry in 1994 and a Ph.D. in Biophysics and Biochemistry in 1999 from The University of Tokyo. He is currently a PI and associate professor in the Department of computational intelligence and systems science at the Tokyo Institute of Technology.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: kiga@dis.titech.ac.jp.

REFERENCES

1 Chin, J. W. Modular approaches to expanding the functions of living matter. *Nat. Chem. Biol.* **2006**, *2*, 304–311.
2 Kaznessis, Y. N. Models for synthetic biology. *BMC Syst. Biol.* **2007**, *1*, 47.
3 Kiga, D.; Sakamoto, K.; Kodama, K.; Kigawa, T.; Matsuda, T.; Yabuki, T.; Shirouzu, M.; Harada, Y.; Nakayama, H.; Takio, K.; Hasegawa, Y.; Endo, Y.; Hirao, I.; Yokoyama, S. An Engineered Escherichia coli Tyrosyl-tRNA Synthetase for Site-Specific Incorporation of an Unnatural Amino Acid into Proteins in Eukaryotic Translation and its Application in a Wheat Germ Cell-Free System. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9715–9720.
4 Soloveichik, D.; Seelig, G.; Winfree, E. DNA as a universal substrate for chemical kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 5393–5398.
5 Gardner, T. S.; Cantor, C. R.; Collins, J. J. Construction of a Genetic Toggle Switch in Escherichia coli. *Nature* **2000**, *403*, 339–342.
6 Elowitz, M. B.; Leibler, S. A Synthetic Oscillatory Network of Transcriptional Regulators. *Nature* **2000**, *403*, 335–338.
7 Seelig, G.; Soloveichik, D.; Zhang, D. Y.; Winfree, E. Enzyme-Free Nucleic Acid Logic Circuits. *Science* **2006**, *314*, 1585–1588.
8 Shlyahovsky, B.; Li, Y.; Lioubashevski, O.; Elbaz, J.; Willner, I. Logic gates and antisense DNA devices operating on a translator nucleic Acid scaffold. *ACS Nano* **2009**, *3*, 1831–1843.
9 Kim, J.; White, K. S.; Winfree, E. Construction of an in Vitro Bistable Circuit From Synthetic Transcriptional Switches. *Mol. Syst. Biol.* **2006**, *2*, 68.
10 Kim, J.; Winfree, E. Synthetic in Vitro Transcriptional Oscillators. *Mol. Syst. Biol.* **2011**, *7*, 465.
11 Montagne, K.; Plasson, R.; Sakai, Y.; Fujii, T.; Rondelez, Y. Programming an in Vitro DNA Oscillator Using a Molecular Networking Strategy. *Mol. Syst. Biol.* **2011**, *7*, 466.
12 You, L.; Cox, R. S., 3rd; Weiss, R.; Arnold, F. H. Programmed Population Control by Cell-Cell Communication and Regulated Killing. *Nature* **2004**, *428*, 868–871.
13 Zhang, D. Y.; Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nat. Chem.* **2011**, *3*, 103–113.
14 Chen, X.; Ellington, A. D. Shaping up nucleic acid computation. *Curr. Opin. Biotechnol.* **2010**, *21*, 392–400.
15 Adleman, L. M. Molecular Computation of Solutions to Combinatorial Problems. *Science* **1994**, *266*, 1021–1024.
16 Takahashi, K.; Yaegashi, S.; Kameda, A.; Hagiya, M. Chain reaction systems based on loop dissociation of DNA. *Lect. Notes Comput. Sci.* **2006**, *3892*, 347–358.
17 Benenson, Y.; Paz-Elizur, T.; Adar, R.; Keinan, E.; Livneh, Z.; Shapiro, E. Programmable and autonomous computing machine made of biomolecules. *Nature* **2001**, *414*, 430–434.
18 Benenson, Y.; Gil, B.; Ben-Dor, U.; Adar, R.; Shapiro, E. An autonomous molecular computer for logical control of gene expression. *Nature* **2004**, *429*, 423–429.
19 Elbaz, J.; Lioubashevski, O.; Wang, F.; Remacle, F.; Levine, R. D.; Willner, I. DNA computing circuits using libraries of DNAzyme subunits. *Nat. Nanotechnol.* **2010**, *5*, 417–422.
20 Torring, T.; Voigt, N. V.; Nangreave, J.; Yan, H.; Gothelf, K. V. DNA origami: a quantum leap for self-assembly of complex structures *Chem. Soc. Rev.* **2011**, DOI: 10.1039/c1cs15057j.

21 Nitta, N.; Suyama, A. Autonomous Biomolecular Computer Modeled after Retroviral Replication. *Lect. Notes Comput. Sci.* **2004**, *2943*, 203–212.

22 Gotoh, O.; Murakami, Y.; Suyama, A. Multiplex cDNA quantification method that facilitates the standardization of gene expression data. *Nucleic Acids Res.* **2011**, *39*, e70.

23 Takinoue, M.; Kiga, D.; Shohda, K.; Suyama, A. Experiments and simulation models of a basic computation element of an autonomous molecular computing system. *Phys. Rev. E: Stat., Nonliner, Soft Matter Phys.* **2008**, *78*, 041921.

24 Sakai, Y.; Mawatari, Y.; Yamasaki, K.; Shohda, K.; Suyama, A. Construction of AND Gate for RTRACS with the Capacityof Extension to NAND Gate. *Lect. Notes Comput. Sc.* **2009**, *5877*, 137–143.

25 Takinoue, M.; Kiga, D.; Shohda, K.; Suyama, A. RNA Oscillator: Limit Cycle Oscillations Based on Artificial Biomolecular Reactions. *New Gener. Comput.* **2009**, *27*, 107–127.

26 Hagiya, M. Perspectives on Molecular Computing. *New Gener. Comput.* **1999**, *17*, 131–151.

27 Guo, P. RNA Nanotechnology: Engineering, Assembly and Applications in Detection, Gene Delivery and Therapy. *J. Nanosci. Nanotechnol.* **2005**, *5*, 1964–1982.

28 Shlyahovsky, B.; Li, D.; Weizmann, Y.; Nowarski, R.; Kotler, M.; Willner, I. Spotlighting of Cocaine by an Autonomous Aptamer-Based Machine. *J. Am. Chem. Soc.* **2007**, *129*, 3814–3815.

29 Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. Cell-Free Translation Reconstituted with Purified Components. *Nat. Biotechnol.* **2001**, *19*, 751–755.

30 , Y.; Kanamori, T.; Ueda, T. Protein synthesis by pure translation systems.Shimizu. *Methods* **2005**, *36*, 299–304.