



Contents lists available at ScienceDirect

Chinese Chemical Letters

journal homepage: www.elsevier.com/locate/ccllet

Allosteric DNzyme-based encoder for molecular information transfer

Minghao Hu^{a,b,1}, Tianci Xie^{a,c,1}, Yuqiang Hu^a, Longjie Li^d, Ting Wang^e, Tongbo Wu^{a,*}^a School of Pharmacy, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China^b Department of Oncology, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China^c GI Cancer Research Institute, Tongji Hospital, Huazhong University of Science and Technology, Wuhan 430030, China^d School of Life Science and Technology, Wuhan Polytechnic University, Wuhan 430023, China^e State Key Laboratory of Bioelectronics, National Demonstration Center for Experimental Biomedical Engineering Education, School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China

ARTICLE INFO

Article history:

Received 27 July 2023

Revised 27 September 2023

Accepted 20 October 2023

Available online 26 October 2023

Keywords:

DNzyme

Encoder

Nucleic acids

DNA circuit

DNA strand displacement

Dynamic DNA nanotechnology

ABSTRACT

Dynamic DNA nanotechnology plays a significant role in nanomedicine and information science due to its high programmability based on Watson-Crick base pairing and nanoscale dimensions. Intelligent DNA machines and networks have been widely used in various fields, including molecular imaging, biosensors, drug delivery, information processing, and logic operations. Encoders serve as crucial components for information compilation and transfer, allowing the conversion of information from diverse application scenarios into a format recognized and applied by DNA circuits. However, there are only a few encoder designs with DNA outputs. Moreover, the molecular priority encoder is hardly designed. In this study, we introduce allosteric DNzyme-based encoders for information transfer. The design of the allosteric domain and the recognition arm allows the input and output to be independent of each other and freely programmable. The pre-packaged mode design achieves uniformity of baseline dynamics and dynamics controllability. We also integrated non-nucleic acid molecules into the encoder through the aptamer design of the allosteric domain. Furthermore, we developed the 2ⁿ-n encoder and the Endo IV-assisted priority encoder inspired by immunoglobulin's molecular structure and effector patterns. To our knowledge, the proposed encoder is the first enzyme-free DNA encoder with DNA output, and the priority encoder is the first molecular priority encoder in the DNA reaction network. Our encoders avoid complex operations on a single molecule, and their simple structure facilitates their application in complex DNA circuits and biological scenarios.

© 2024 Published by Elsevier B.V. on behalf of Chinese Chemical Society and Institute of Materia Medica, Chinese Academy of Medical Sciences.

DNA nanotechnology based on nucleic acid materials is highly programmable and functionally diverse, showing its application potential in various fields such as information processing [1,2], logic operations [2,3], dynamic network construction [4,5], molecular sensing [6,7], intelligent diagnosis of diseases [8], and intelligent therapy [9]. Dynamic DNA nanotechnology is founded on the basic principle of Watson-Crick base pairing to construct non-equilibrium nucleic acid networks and nanomachines. Toehold-mediated strand displacement (TMSD) serves as the backbone of dynamic DNA nanomachines and nanonetworks. Therefore, the development and modification of strand displacement tools became an important element of DNA nanotechnology, and researchers are actively working on developing new strand displacement tools to drive the reaction of DNA [10,11]. However, strand displacement

reaction can lead to crosstalk due to the presence of repetitive sequences in the branch migration domain, thus limiting its application. As a consequence, the strategy of relying exclusively on strand displacement encounters challenges in the construction of some complex circuit elements, such as the encoder.

The encoder is an essential universal element for information conversion and compilation. It converts information from upstream pathways into binary DNA codes that can be utilized in subsequent DNA circuits and nucleic acid calculations [12]. In precision cancer diagnosis and treatment, nucleic acid computation based on tumor microenvironment and biomarkers is the core of intelligence [8,9]. Standardized processing and signaling of DNA information are necessary for articulating different DNA circuits. Encoders can convert DNA information for transmission to downstream DNA circuits to perform nucleic acid calculations in these application scenarios. However, the current development of molecular encoders is limited and often uses fluorescent or chemical signals as output [13–19], making it challenging to perform the

* Corresponding author.

E-mail address: wutongbo@hust.edu.cn (T. Wu).¹ These authors contributed equally to this work.

information transfer role of encoders. These challenges may be attributed to the difficulty of development posed by sequence limitations of strand displacement and unwanted reactions. Previously our team built encoders using the Exponential Amplification Reaction (EXPAR) of polymerase and nickase, but the introduction of enzymes limited their application in some scenarios [12]. Moreover, integrating non-nucleic acid molecules into nucleic acid calculations is an important task of encoders, but it is rarely implemented. Several non-nucleic acid molecule integration approaches based on strand displacement reactions have been proposed, but they are mostly sequence-dependent and lack robustness and dynamic controllability [20,21].

DNAzyme is widely used in biomarker detection [6,22-26], molecular machines [25,27-29], DNA circuits [5,30-35], and cancer therapy [36,37] owing to its unique dual role as both DNA and enzyme [37]. However, despite its successful application in DNA circuits, the complex design, poor scalability, and limited control over dynamics have hindered its further patterned applications. To overcome these challenges, our team has proposed an allosteric DNAzyme (ANAZyme) that extends the stimulus-response toolbox and lays the foundation for molecular encoder construction [26].

In this study, we report a scalable ANAZyme-based encoder for information transfer. To achieve consistent baseline dynamics and dynamic regulation by a regulator, we use pre-packaged ANAZyme where the ANAZyme is pre-bound to the substrate strand. The allosteric domain of the ANAZyme serves as the input response, while the substrate strand acts as the output. Importantly, the input and output sequences are freely programmable and unrelated to each other, enabling versatile applications. Our encoder demonstrates the ability to handle different input forms, including DNA and non-nucleic acid molecules. And it is independent of dynamics and strand displacement sequences. Moreover, we developed scalable 2^n - n encoders and the Endo IV-assisted priority encoder. Based on this information processing model, our encoder's input and output forms are simpler, making it easier to use in other applications.

Fig. 1A shows the basic structure of the ANAZyme. When there is an upstream DNA input, the allosteric domain binds to the DNA, and the core sequence of ANAZyme is clustered and becomes active, thus cleaving the ribonucleic acid sites in the substrate. We first investigated three reaction modes for ANAZyme to activate downstream reactions: pre-packaged mode, double-stranded DNA (dsDNA) block mode, and hairpin block mode. In the pre-packaged mode, the ANAZyme and substrate strands are annealed at a 1:1 ratio, forming a stable pre-bound structure that does not require recognition. The cleavage site in the substrate is the RNA base adenine (rA, red dot in Fig. 1) for the ANAZyme. The 3'-end from the rA of the substrate strand has a nine-nucleotide (9-nt) complementary sequence to ANAZyme, which is insufficient to form a stable dsDNA structure. Thus, after the substrate strand is cleaved, the part of the 3'-end falls off to form the output and react with the downstream through the toehold domain (purple in Fig. 1). However, without cleavage, the ANAZyme and the substrate strand can form a stable pre-packaged structure to avoid reacting with the downstream because the 5'-end from the rA also has a 10-nt complementary sequence. In the dsDNA block mode, the sequence near the substrate strand's rA forms a bubble structure that serves as a recognition domain for ANAZyme. In the hairpin block mode, the toehold domain and a portion of the migration domain in the substrate strand are blocked by the stem of the hairpin. When the ANAZyme cleaves the substrate strand, the hairpin breaks down into two parts to form the output. The same ANAZyme sequence (8-17E) was used to evaluate these three modes. The concentrations of ANAZyme, substrate strands, and reporter were identical.

To investigate the factors affecting the performance of the three modes, an optimization study was carried out for each of them

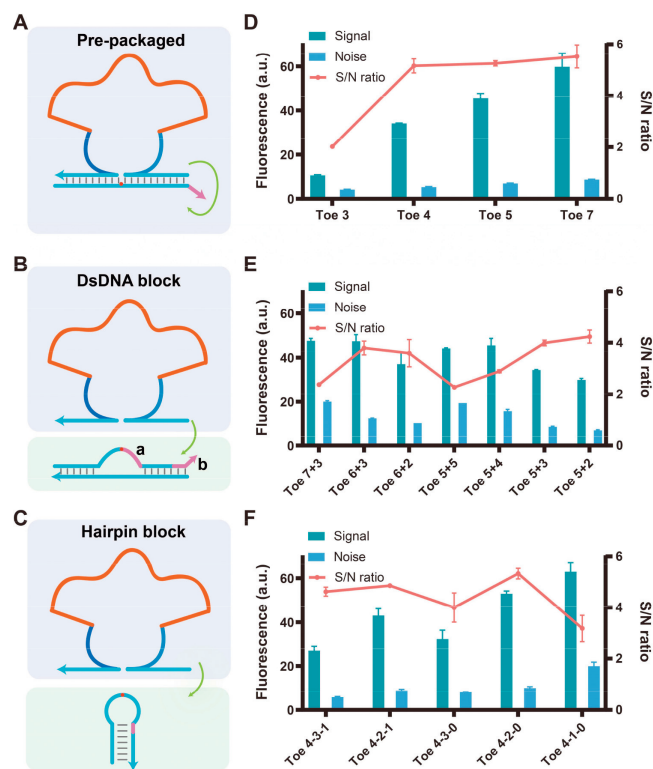


Fig. 1. Reaction mode of Encoder. (A) Schematic diagram of the pre-packaged mode. (B) Schematic diagram of double-stranded block mode. (C) Schematic diagram of hairpin block mode. (D-F) Optimization study of the three reaction modes. The length of each part of the substrate strand was optimized. Signal values were intercepted at the point 20 min after the reaction. The toe ($a+b$) in (B) denotes the length of the (ring + 3'-end) toehold. In toe ($a-b-c$) in (C), a denotes the length of the forward toehold, b denotes the length of the reverse toehold, and c denotes the additional complementary length of the stem. Detailed diagrams are shown in Figs. S5 and S7 (Supporting information).

(Figs. 1D-F). In the pre-packaged mode, the ANAZyme and the substrate have 19 nt complementary base pairs (9 nt + 10 nt), which results in an excellent blocking effect, and the noise in Fig. 1D was small. However, the conventional split DNAzyme does not produce a good blocking effect (Fig. S1 in Supporting information), indicating the advantage of ANAZyme. We also evaluated the effect of the binding length between the allosteric domain and the input strand on the reaction. When 12 bases of the allosteric domain are complementary to the input strand, a stable structure cannot be formed, and the ANAZyme is difficult to be activated. While the length of 16 bases has a good activation effect (Fig. S2 in Supporting information). In addition, the influence of the allosteric domain's length on the blocking effect was not significant (Fig. S3 in Supporting information). We explored the output signals induced by different input concentrations (Fig. S4 in Supporting information). In pre-packaged mode, ANAZyme is unable to achieve free circular cleavage. The concentration of the input is related to the extent and rate of the output.

The dsDNA block mode has been used to construct logic gates [35]. The toehold domain in the substrate strand consists of two parts (toe a and b , purple in Fig. 1B). It is worth noting that only the toehold at the 3'-end is the true toehold (forward toehold, the corresponding part of the reporter is exposed), while the sequence of the reporter corresponding to the loop part (reverse toehold) is blocked (detailed in Fig. S5 in Supporting information). As shown in Fig. 1E, the dsDNA block mode has relatively high leakage and a slightly lower S/N ratio ($S/N < 4$), although the structure leakage is lower in the original design [35]. This hints at the poor robustness of this structure. Interestingly, when we used the optimized

structure in ANAzyme 1 (8–17E) for ANAzyme 2 (10–23), the signals of Off and On became indistinguishable (Fig. S6 in Supporting information). Although the toehold, loop, and stem have the same length, differences in sequence and ANAzyme may disrupt this delicate structure.

The optimization of the hairpin mode is shown in Fig. 1F. In the “toe a-b-c” in Fig. 1F, “a” denotes the length of the forward toehold, “b” denotes the length of the reverse toehold, and “c” denotes the additional complementary length of the stem. The detailed diagram is shown in Fig. S7 (Supporting information). In verifying the hairpin block effect, we identified that the independent forward toehold caused leakage (Fig. S8 in Supporting information). Therefore, we solved this problem by reducing the length of the substrate strand and adding the reverse toehold (Figs. S7 and S8 in Supporting information).

Although the optimization study of the three modes was explored without rigorous comparisons and calculations, the pre-packaged and hairpin block modes may be the more robust reaction modes. To further validate the generality of the pre-packaged mode, we evaluated the binding length of the output domain and substrate and the effects and sequences (Figs. S9 and S10 in Supporting information). When the binding length is greater than or equal to 17 nt, ANAzyme has a better blocking effect on the substrate strand (Fig. S9). Changes in the sequence of the allosteric domain and the output domain did not have a significant effect on either ANAzyme (Fig. S10). In addition, the pre-packaged mode shows faster reaction dynamics (Fig. S11 in Supporting informa-

tion), likely due to the pre-bound structure not requiring a recognition and binding process.

As an element for information conversion and transfer, the encoder should have controllable dynamic reactions to meet the needs of different downstream nucleic acid reaction networks. To achieve this, we designed a dual regulation mechanism that allows for regulation of the reaction dynamics by toehold length and spacer length, respectively (Figs. 2A and B). The toehold length is responsible for coarse regulation, while the spacer length is responsible for fine regulation. In Figs. 2C and D, S_nT_m denotes the output with an n -nt spacer and an m -nt toehold. We first verified this reaction process with native-PAGE (Fig. S12 in Supporting information). As shown in Fig. 2C, a change of only two bases in the toehold length can cause a dramatic change in the reaction dynamics. The reaction rate is low at a toehold length of 7 nt in Fig. 2C, while a 7-nt toehold without a spacer can cause a significant rate increase in Fig. 1E. This contradiction may be caused by the instability of the dsDNA introduced by the spacer. Changes in the spacer length can bring about slight changes in reaction dynamics and plateau values (Fig. 2D). This dual regulatory mechanism can provide combinations for dynamics regulation freely. It is achieved primarily through the structure of the output rather than the activity of the ANAzyme, which preserves a concise structure of the ANAzyme and the allosteric domain and allows for increased scalability.

Negative regulation mode is vital in both organisms and artificial networks. The annihilation gate is one of the most common

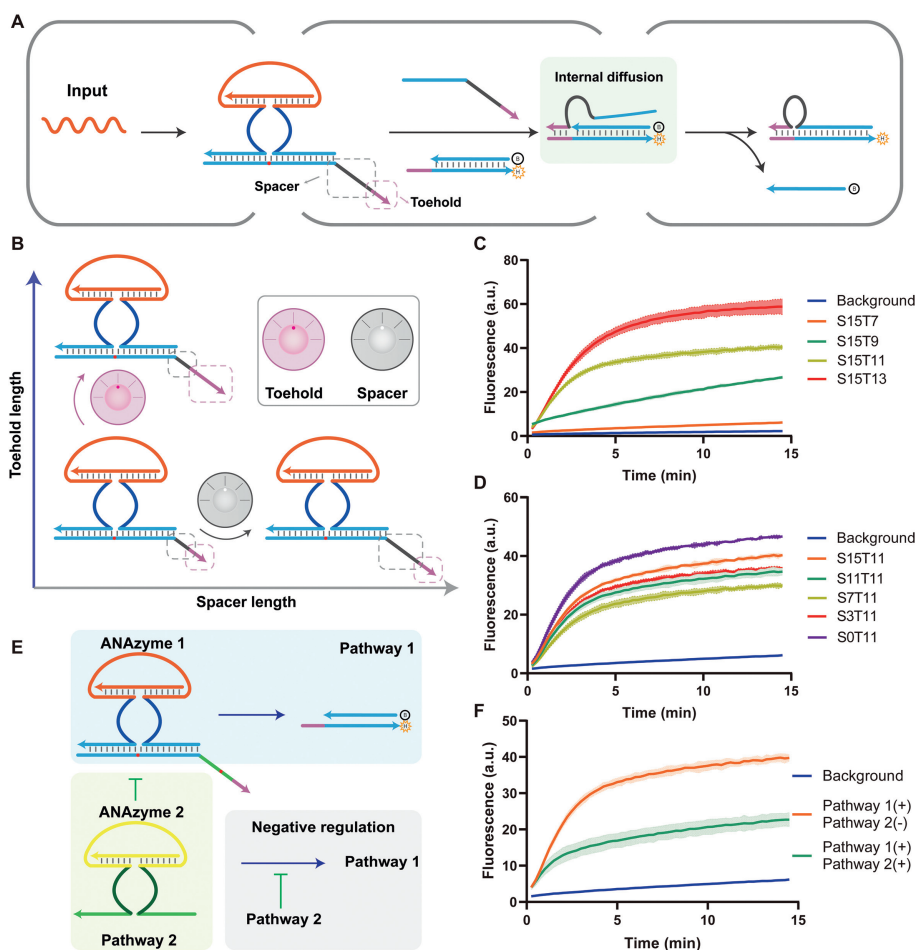


Fig. 2. Regulation of the reaction dynamics of the encoder. (A) Reaction process with the introduction of a regulatory mechanism. (B) Dual regulation mechanism of encoder reaction dynamics. Spacer and toehold can synergistically regulate encoder reaction dynamics. (C) The fluorescence profile of different spacer lengths with the same toehold length. Spacer has a fine-tuning function. (D) The fluorescence profile of different toehold lengths with the same spacer length. Toehold has a coarse tuning function. (E) Negative regulation mode of the encoder. Pathway 2 activation causes an inhibitory effect on Pathway 1. (F) The fluorescence profile of the negative regulation mode.

types of negative regulation, where the target strand is eliminated by a blocking strand that is complementary to it. However, the high sequence match between the annihilation gate and the target strand limits its application in DNA circuits, and the controlled release of negative regulation is complex. We developed a new pathway regulation by introducing ANAzyme 2's substrate sequence in the spacer domain of ANAzyme 1's substrate (green in Fig. 2E). When activated, ANAzyme 2 cleaves the spacer's ribonucleic acid sites (red dot in Fig. 2E), separating the toehold from the migration domain. As a result, the substrate strand loses its ability to carry out strand displacement with the downstream reporter, and pathway 1 is inhibited. As shown in Fig. 2E, the signal is significantly reduced when both pathways are activated simultaneously (green curve in Fig. 2F) compared to only pathway 1 is activated (orange curve in Fig. 2F). Unlike common DNA annihilation gates, our pathway inhibition is carried out by ANAzyme cleavage rather than the blocking effect of the complementary strand. This design allows for reduced sequence interference generated by the annihilation gate. Furthermore, this ANAzyme-mediated pathway inhibition is characterized by a controllable release, further enhancing its flexibility in DNA circuits.

Although nucleic acid computing can solve informational and biomedical problems, the gap in integrating non-nucleic acid molecules limits its application. Thus, the development of molecular encoders for non-nucleic acid inputs to cope with different environments is necessary. Aptamers provide a powerful recognition tool for different inputs, such as small molecules and proteins [21,38], and are widely used for integrating non-nucleic acid molecules into DNA circuits. However, the competing model used in integrating non-nucleic acid molecules into DNA circuits depends on the aptamer sequence on the target strand sequence, limiting its usability [21]. Zhang *et al.* developed a generic non-nucleic acid integration method based on the remote toehold-mediated strand displacement model, using spacer domain chimeras to solve this problem [20]. However, the initial study of remote toehold and our previous study showed that, in some cases, remote toehold-mediated strand displacement reactions could occur even if the spacer is in an uncompact structure [39,40]. This may imply that the non-nucleic acid integration strategy based on remote toehold has an entropy penalty limitation and is more demanding on the sequence. We used an allosteric DNAzyme design [26] to integrate the aptamer into the allosteric domain (red part in Fig. 3A). The integration strategy based on allosteric DNAzyme can avoid the above problems. On the one hand, the separation of input and output guarantees that the output sequence is independent of the aptamer sequence. On the other hand, the combination of ligand and aptamer does not affect the reaction dynamics. The allosteric domain can respond to both the DNA input and the ligand. Here, we use ATP as a proof of concept. When ATP is introduced, ATP binds to the aptamer of the allosteric domain (red part in Fig. 3A). Uniting the complementary 5 bases on the other side, the allosteric domain aggregates, bringing the core region of the DNAzyme in close proximity and generating activity. When DNA is input, the DNA binds directly to the allosteric domain, producing an active DNAzyme structure. The fluorescence curve in Fig. 3B demonstrates the ability of the encoder to respond to both DNA input and ATP input. Both DNA and ATP activate the encoder well and produce similar dynamic properties. This illustrates that our encoder has good dynamics controllability without significant dynamics change and entropy penalty due to aptamer introduction, conferred by the robustness of the DNAzyme-based encoder.

In upstream information processing, the same information often needs to be used as input to different pathways or to perform different functions. Designs dominated entirely by strand displacement may create problems with a high correlation of different pathways, resulting in crosstalk. It can also lead to disconnec-

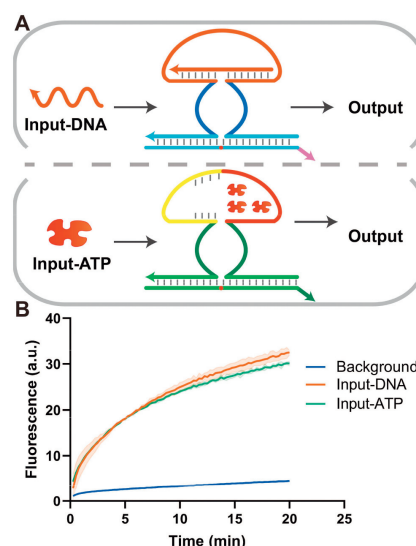


Fig. 3. Integration of non-nucleic acid molecules. (A) Schematic diagram of molecular input. ATP can activate the DNAzyme by binding to the aptamer, which compacts the allosteric domain. Five complementary paired bases on the other side of the aptamer sequence are used for auxiliary binding. The DNA input can also bind to the allosteric domain. (B) The fluorescence profile of the DNA input and the molecular input. Both inputs can activate the encoder. The reaction was characterized using a downstream reporter.

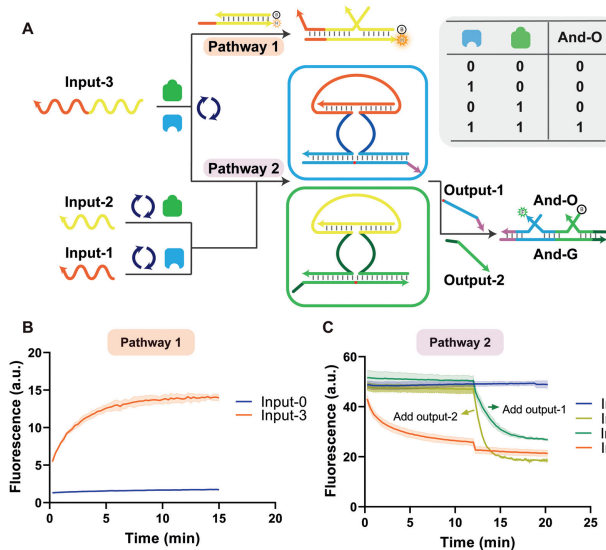


Fig. 4. Pathway conversion function of the encoder. (A) Schematic diagram of the encoder performing pathway conversion. (B) The fluorescence profile of input-3 performing the pathway 1 task. (C) The fluorescence profile of the AND gate of pathway 2.

tions between the upstream signal and the downstream response when the downstream response of a pathway requires a specific sequence. Here, we demonstrate the pathway conversion function of the encoder. The same input, input-3, can perform the task of pathway 1 and pathway 2 *via* encoder conversion (Fig. 4A). Input-3 was initially tasked with performing a one-step strand displacement reaction in pathway 1 (Fig. 4A). In pathway 2, we set up an AND gate assignment to be performed. We use two different DNAzyme (DNAzyme 1 and DNAzyme 2) as the upstream interfaces to the AND gate. DNAzyme 1 is pre-packaged with output-1, and DNAzyme 2 is pre-packaged with output-2. When we need input-3 to perform the assignment of pathway 2, we can add the encoder node. Input-3 can bind to DNAzyme 1 and DNAzyme 2

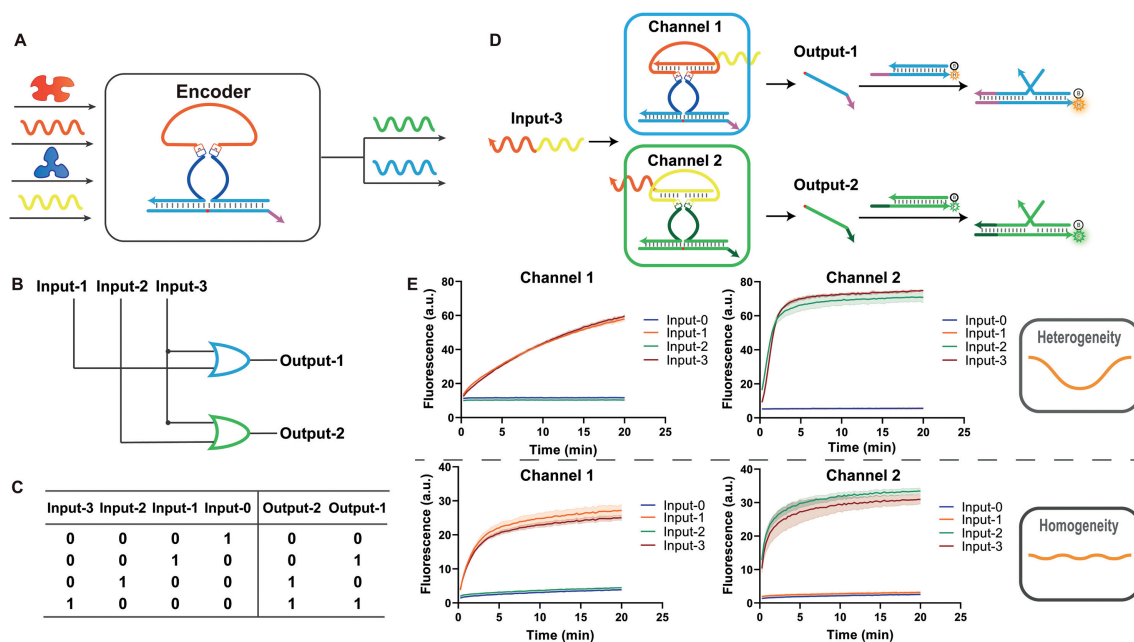


Fig. 5. Construction of the 4-2 encoder. (A) Schematic diagram of the 4-2 encoder. The encoder can respond to four different inputs, encode them, and produce a combination of two outputs. (B) Circuit diagram of the 4-2 encoder. Each channel performs AND gate operations on the inputs. (C) Truth table of the 4-2 encoder. (D) The process of the 4-2 encoder responding to input-3. The variational region of the encoder allows recognition of the input. (E) The fluorescence profiles of the 4-2 encoder in operation. Channel 1 and channel 2 are characterized by HEX and FAM, respectively. The top plots do not use the pre-packaged mode and are characterized by the substrate strand labeled with fluorescent and quench groups at each end. The bottom plots use the pre-packaged mode and are characterized by the downstream dsDNA reporter. The different modes show heterogeneous and homogeneous dynamics patterns.

allosteric domains to generate output-1 and output-2. The reaction with the AND gate can only occur when both outputs are present simultaneously. We characterized pathway 1 and pathway 2 using two different fluorescent groups, hexachloro-fluorescein (HEX) and fluorescein (FAM), respectively. The fluorescence profile shows input-3 performing the task of pathway 1 (Fig. 4B). Meanwhile, input 3 reacts with two DNAzymes in pathway 2 to produce output-1 and output-2. Two outputs react with AND gate to displace the AND-Output (AND-O) strand (Fig. 4C). 5'- and 3'-ends of the AND-O strand are labeled with a fluorophore (FAM) and a quencher (BHQ), respectively. When AND-O is assembled with AND-Gate (AND-G), FAM, and BHQ are far away, and FAM can generate a fluorescent signal. When AND-O is replaced, AND-O exists as a single strand, and coiling occurs. FAM and BHQ become close, and the fluorescence is quenched due to fluorescent resonance energy transfer (FRET). Input-1 and input-2 perform the tasks of pathway 2 through the encoder, but neither can activate the AND gate (Fig. 4C). To verify that input-1 and input-2 were still functional, we added output-2 to the system with input-1 input and output-1 to the system with input-2 after the reaction started 12 min later. The AND gate was activated when output-2 was added (yellow curve in Fig. 4C). The same phenomenon occurs for the input-2 system.

With the encoder's pathway conversion function, the same input can perform different tasks depending on the need, and these tasks are entirely independent. In addition, the encoder can be implemented with a single input AND gate design.

The encoder has up to 2^n inputs and n outputs, and the 2^n inputs are encoded in binary as different combinations of the n outputs. Fig. 5A shows the design concept of the 4-2 encoder. The encoder can respond to 2^2 DNA strands or non-nucleic acid molecules and outputs two standard DNA signals for downstream design. Figs. 5B and C show the circuit diagram and truth table of the 4-2 encoder. The circuit diagram (Fig. 5B) shows that the encoder performs the OR gate function. Each channel corresponds to an output, and each can perform an OR gate on all inputs. However, the operation of the OR gate requires a complex design,

which limits the design and application of encoders and may be one of the reasons why molecular encoders are challenging to develop.

The structures of the F_c segment of immunoglobulins (IgG, for example) are all highly conserved and belong to the Constant region. At the same time, the N-terminal of the corresponding F_{ab} , which has a very high amino acid sequence diversity, is termed the Variable region. This coincides with their functions, as the sequence of the complementarity determining regions (CDR) in the F_{ab} segment is associated with antibody-specific recognition of antigens. The F_c segment is where immunoglobulin interacts with effector molecules or cells. This structure may be a reason for the limited effector cell/molecule processing of variable antigens *in vivo*. We found similarities between this molecular processing pattern and the ANAzyme (Fig. S13 in Supporting information). Here, we take advantage of the separate and reprogrammable input and output properties of the ANAzyme to design different channels for the combination of substrate strands, the recognition arm of the ANAzyme (which binds to the substrate strand), and the catalytic core of the ANAzyme. The response to input by the allosteric domain can be separated from the recognition arm, and the same channel can correspond to different allosteric domains of input response. Fig. 5D shows our design, where both channel 1 and channel 2 can have a single response to input-3 and generate a signal. When we need channel 1 to respond to a distinct input 1, we can individually reprogram the sequence of allosteric domains of channel 1 (Fig. S14 in Supporting information). Each channel in the system will be associated with a number of different allosteric domains. This encoder design is scalable, and the 2^n-n design is based on the same principle.

We demonstrate our design with a 4-2 encoder. The output results are consistent with the truth table when the encoder responds to different inputs, and the pre-packaged mode demonstrates good leakage suppression (Fig. 5E). In addition, our encoder can be normalized to the dynamics. The fluorescence profile at the top of Fig. 5E shows the results without the pre-packaged

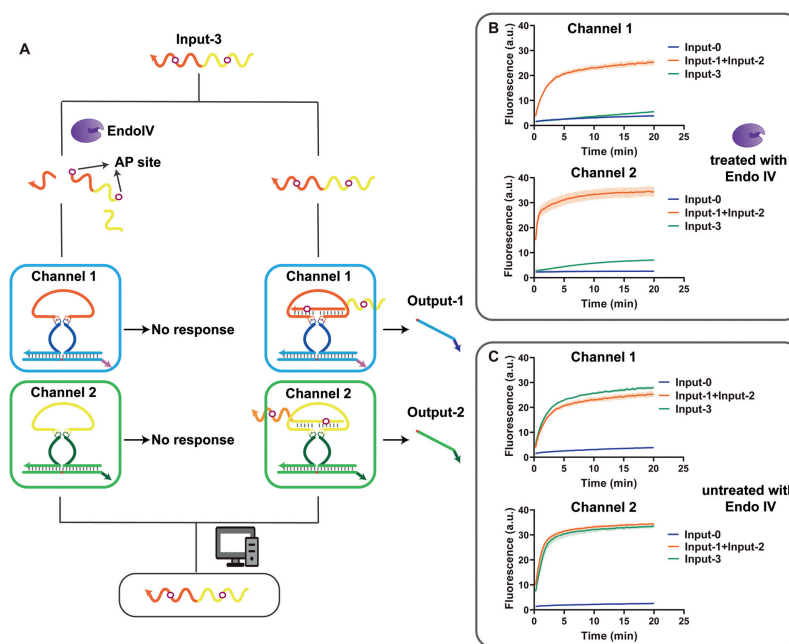


Fig. 6. Endo IV-assisted priority encoder. (A) flowchart of the priority encoder. Parallel reactions are performed on the same input, one reaction is treated with Endo IV, and the other is not. The input is differentiated by the output signal. (B) The fluorescence profile of the priority encoder with Endo IV treatment. (C) The fluorescence profile of the priority encoder without Endo IV treatment.

mode, characterized using a substrate strand labeled with a fluorophore and a quenching group. The fluorescence profile at the bottom of Fig. 5E shows the results with the pre-packaged mode, characterized using a downstream dsDNA reporter. In the absence of the pre-packaged mode, heterogeneous dynamics effects were generated due to the difference in DNase, although the binding length of the recognition arm and substrate strand is the same in both DNases. A homogeneous dynamic effect is demonstrated in the pre-packaged mode because the DNase and substrate recognition and the binding process have been completed. The pre-packaged mode can help to construct the standardized output of the encoder.

The encoder can only perform operations on one input at a time while producing incorrect results when multiple inputs are introduced simultaneously. For example, when input-1 and input-2 are present at the same time, both channel 1 and channel 2 are activated. In the presence of input-3, both channel 1 and channel 2 can also be activated simultaneously, making it impossible to distinguish between input-1 + input-2 and input-3. The priority encoder was developed to solve this problem. Each input has a different rank in a priority encoder for electronic circuits. When multiple inputs are present, the priority encoder only calculates the highest-ranked input and ignores the lower ones. Fig. S15 (Supporting information) shows the truth table for a 4-2 priority encoder. In electronic circuits, the priority encoder needs to be implemented through a complex collection of logic gates [41], which is challenging to implement in the field of DNA circuits. Therefore, to our knowledge, there is no molecular-level priority encoder. Even then, priority encoders can only output to the highest level input at the same time and cannot perform parallel operations. However, without parallel operations, the transmission efficiency of the DNA circuit would be significantly reduced. Moreover, the input information is artificially graded, and ignoring lower-grade information may, in many cases, be unacceptable. Therefore, we wish to develop priority encoders that can distinguish between different inputs while retaining parallel operations of encoders.

We developed an Endo IV-assisted 4-2 priority encoder. Endo IV is a nucleic acid endonuclease that recognizes apurinic/

aprimidinic (AP) sites of dsDNA and cleaves the phosphodiester bond 5' to the lesion generating a hydroxyl group at the 3'-terminus [42,43]. The same input performs two separate reactions simultaneously, one is pre-treated with Endo IV before the reaction, and the other remains untreated (Fig. 6A). We modified two AP sites on input-3. When input-3 is treated with Endo IV, input-3 will break into three parts and cannot bind to ANAzyme to generate the signal. While input-3 without Endo IV treatment would produce a normal signal. Input-1 and input-2 are not modified with the AP site and produce the same signals with or without Endo IV treatment. The fluorescence reading of the output allows us to distinguish between input-1 + input-2 and input-3. At this point, the Endo IV produces good cleavage activity. Figs. 6B and C show the complete input, output, and read processes in one run. When input-3 passes the checkpoint of Endo IV, the signal of input-3 is drastically reduced (Fig. 6B), while the fluorescence rate of the untreated group is speedy (Fig. 6C).

This Endo IV-assisted information processing mode is simple and easy to implement without complex calculations, and the input information can be retained. The priority encoder of this mode can be extended to 7-3 carry bits (Fig. S16 in Supporting information), but further extensions are still challenging.

In this study, we reported a scalable encoder based on ANAzyme for information transfer. We first explored the reaction mode of the encoder. Although the double-stranded block mode [35] and the hairpin block mode [44] already exist, studies have yet to compare them directly. We found that the delicate structure of the dsDNA block mode may suggest poorer robustness, while the pre-packaged mode has better dynamic controllability. The dual regulation consisted of coarse and fine regulation controlled by the toehold and the spacer, respectively. Integrating non-nucleic acid molecules helps DNA machines and networks face real and diverse application scenarios. We use aptamer tools to integrate non-nucleic acid molecules and use ATP as a proof of concept. Taking the above advantages, we proposed a scalable 2^n - n encoding scheme with the ANAzyme. Finally, we introduced Endo IV and parallel reactions to solve the problem of priority encoders. This model is simple, fault-tolerant, and retains the ability to

compute in parallel, although it suffers from the limitations of scalability and simultaneous processing of three inputs.

The ANAzyme-based encoder enables the encoding process to be carried out entirely by DNA, avoiding the introduction of enzymes. The input and output sides of the DNAzyme are freely programmable, and the allosteric domain and output sides are independent of each other. Inspired by immunoglobulins' molecular structure and effect pattern, we have preserved the structural simplicity of individual ANAzyme and avoided complex design. These properties ensure the simplicity of the encoder, making it easier to integrate into complex DNA networks and reducing the constraints imposed by complex secondary design. The proposed ANAzyme-based encoder may provide more practical concepts and strategies in DNA nanotechnology, leading to the development of novel nanomachines, DNA networks, and nanomedicine.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (No. 82172372), and the Opening Research Fund of State Key Laboratory of Digital Medical Engineering (No. 2023-M04).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ccl.2023.109232.

References

- [1] D. Huang, H.Y. Han, C. Guo, et al., *Nanoscale* 13 (2021) 5706–5713.
- [2] Q. Ma, C. Zhang, M.Z. Zhang, D. Han, W.H. Tan, *Small Struct.* 2 (2021) 2100051.
- [3] C.Y. Zhou, Y.W. Song, X.Y. Jin, B. Li, C.Y. Pang, *Nanoscale Horiz.* 8 (2023) 176–184.
- [4] L. Yue, S. Wang, Z.X. Zhou, I. Willner, *J. Am. Chem. Soc.* 142 (2020) 21577–21594.
- [5] Z.X. Zhou, J.B. Wang, R.D. Levine, F. Remacle, I. Willner, *Chem. Sci.* 12 (2021) 5473–5483.
- [6] Y.R. Cao, H.Q. Zhang, X.C. Le, *Anal. Chem.* 93 (2021) 15712–15719.
- [7] S.M. Traynor, G.A. Wang, R. Pandey, F. Li, L. Soleymani, *Angew. Chem. Int. Ed.* 59 (2020) 22617–22622.
- [8] Q.N. Kou, L. Wang, L.H. Zhang, et al., *Small* 18 (2022) 2205191.
- [9] J. Chen, S.N. Fu, C.Y. Zhang, H.Y. Liu, X. Su, *Small* 18 (2022) 2108008.
- [10] Y.W. Liao, H. Hu, X.F. Tang, et al., *Nucl. Acids Res.* 51 (2023) 29–40.
- [11] X.L. Yang, Y.N. Tang, S.M. Traynor, F. Li, *J. Am. Chem. Soc.* 138 (2016) 14076–14082.
- [12] T.C. Xie, Y.H. Deng, J.R. Zhang, et al., *Nucl. Acids Res.* 50 (2022) 8431–8440.
- [13] D.Q. Fan, J.B. Zhu, Y.Q. Liu, E.K. Wang, S.J. Dong, *Nanoscale* 8 (2016) 3834–3840.
- [14] Y. He, Y.T. Chen, C.Y. Li, H. Cui, *Chem. Commun.* 50 (2014) 7994–7997.
- [15] D. Kang, R.J. White, F. Xia, et al., *Npg Asia Mater* 4 (2012) e1.
- [16] S. Liu, M.L. Li, X. Yu, C.Z. Li, H.Y. Liu, *Chem. Commun.* 51 (2015) 13185–13188.
- [17] T. Nishimura, Y. Ogura, K. Yamada, Y. Ohno, J. Tanida, *Biomed. Opt. Express* 5 (2014) 2082–2090.
- [18] C.T. Wu, C.Y. Zhou, E.K. Wang, S.J. Dong, *Nanoscale* 8 (2016) 14243–14249.
- [19] H.L. Li, Y.Q. Liu, S.J. Dong, E.K. Wang, *Npg Asia Mater.* 7 (2015) e166.
- [20] Q.L. Zhang, Y. Wang, L.L. Wang, et al., *Angew. Chem. Int. Ed.* 62 (2023) e202214698.
- [21] J.D. Munzar, A. Ng, D. Juncker, *Chem. Soc. Rev.* 48 (2019) 1390–1419.
- [22] H.Y. Peng, X.F. Li, H.Q. Zhang, X.C. Le, *Nat. Commun.* 8 (2017) 14378.
- [23] L.Y. Duan, J.W. Liu, R.Q. Yu, J.H. Jiang, *Biosens. Bioelectron.* 177 (2021) 112976.
- [24] S. Khan, B. Burciu, C.D.M. Filipe, et al., *ACS Nano* 15 (2021) 13943–13969.
- [25] Y.J. Zhao, Y. Liao, J.H. Fu, et al., *Biosens. Bioelectron.* 219 (2023) 114757.
- [26] Y.Q. Hu, C.J. Li, M.H. Hu, et al., *Small* 19 (2023) 2300207.
- [27] W.C. Deng, J.Y. Xu, H.Y. Peng, et al., *Biosens. Bioelectron.* 217 (2022) 114704.
- [28] S.D. Mason, G.A. Wang, P. Yang, Y.Y. Li, F. Li, *ACS Nano* 13 (2019) 8106–8113.
- [29] Y.Q. Hu, Z. Zhang, W. Zhang, et al., *Chin. Chem. Lett.* 33 (2022) 3026–3030.
- [30] G. Shi, C. Yan, J.H. Chen, *Anal. Chem.* 93 (2021) 3273–3279.
- [31] J. Wei, H.M. Wang, Q. Wu, et al., *Angew. Chem. Int. Edit.* 59 (2020) 5965–5971.
- [32] C. Zhang, J. Yang, S.X. Jiang, Y. Liu, H. Yan, *Nano Lett.* 16 (2016) 736–741.
- [33] X.D. Zheng, J. Yang, C.J. Zhou, et al., *Nucleic Acids Res.* 47 (2019) 1097–1109.
- [34] J. Yang, R.F. Wu, Y.F. Li, et al., *Nucleic Acids Res.* 46 (2018) 8532–8541.
- [35] J. Elbaz, O. Lioubashevski, F.A. Wang, et al., *Nat. Nanotechnol.* 5 (2010) 417–422.
- [36] Z. Wang, J. Yang, G. Qin, et al., *Angew. Chem. Int. Ed.* 61 (2022) e202204291.
- [37] J. Jouha, H. Xiong, *Small* 17 (2021) 2105439.
- [38] G.A. Wang, X.H. Wu, F.F. Chen, et al., *J. Am. Chem. Soc.* 145 (2023) 2750–2753.
- [39] M. Hu, Z. Chu, H. Wang, W. Zhao, T. Wu, *Microchem. J.* 190 (2023) 108706.
- [40] A.J. Genot, D.Y. Zhang, J. Bath, A.J. Turberfield, *J. Am. Chem. Soc.* 133 (2011) 2177–2182.
- [41] S. Abdel-Hafeez, S. Harb, *IEEE Trans. Circuits Syst. II Express Briefs* 53 (2006) 597–601.
- [42] W. Yuan, Z. Zhang, Y. Hu, et al., *Sens. Actuators B: Chem.* 359 (2022) 131575.
- [43] Z. Zhang, Y. Hu, W. Yuan, Y. Deng, T. Wu, *Sens. Actuators B: Chem.* 357 (2022) 131378.
- [44] M. Hu, Y. Hu, T. Wu, *Talanta* 259 (2023) 124546.