Enzyme-Free, Isothermal Noncovalent DNA Catalytic Reactions

by

Jing Yang Xu

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Laboratory Medicine and Pathology

University of Alberta

Abstract

MicroRNA is a class of endogenous, noncoding, short RNA molecules and the aberrant expression of microRNA is related to the development of a variety of human diseases. The detection of microRNA in living cells is vital to understanding its dynamic functions and profiles, diagnosing diseases, and developing microRNA-based therapy. My thesis research focuses on the development of enzyme-free, isothermal amplification techniques applicable to the detection of nucleic acids such as microRNA.

Non-covalent DNA catalytic reactions represent a recent major advance of dynamic DNA nanotechnology. These reactions can produce multiple output DNA strands from a single nucleic acid input through cycling toehold-mediated strand displacement and/or toehold exchange reactions. The programmability and the signal amplification capability of these catalytic reactions have been utilized for building DNA nanomachines, realizing molecular computation, and detecting disease biomarkers. These DNA catalytic reactions can achieve amplified detection of nucleic acids without the need of any enzymes and washing steps at room temperature and 37 °C.

Three main DNA catalytic reactions have been reported previously, including hybridization chain reaction (HCR), catalytic hairpin assembly (CHA), and entropy-driven DNA catalysis (EDC). However, the application of these reactions is limited by their complexity in design and low amplification efficiency. The primary objective of my thesis research was to develop improved DNA catalytic reactions by destabilizing the substrate of the toehold exchange catalytic reaction to make the forward reaction energetically favorable. The toehold exchange catalytic reaction, requiring only a dsDNA substrate with a toehold and a ssDNA fuel, is chosen for its simplicity. I hypothesize that the substrate can be destabilized by introducing DNA

structures, mismatches or chemically modified nucleotides into the substrate. Modification of the substrate serves several purposes. First, it raises the free energy of hybridization of the substrate ($\Delta G^o_{f_{Substrate}}$) making the overall reaction energetically favorable, ΔG^o_{OR}<0. Secondly, raising the $\Delta G^o_{f_S}$ makes the formation of intermediates more energetically favorable. Finally, the modification of the substrate can affect the kinetics of the reverse reaction to favor the formation of products.

I developed a three-way-junction (TWJ) mediated DNA catalytic reaction. Without the TWJ, the products of the toehold exchange catalytic reaction is identical to the reactants, and the reaction free energy change, ΔG^o_{OR}, is zero. The TWJ disrupts the base-stacking interaction and destabilizes the substrate, making the ΔG^o_{OR} negative. Compared to without TWJ, the introduction of TWJ improved amplification efficiency of the DNA catalytic reaction by approximately 13-fold within one hour. To further improve the amplification efficiency, we cascaded two TWJ-mediated DNA catalytic reactions, where the output DNA of the first reaction acted as the input DNA for the second reaction. This two-layer reaction technique afforded >250-fold amplification in 1 hour and lowered the limit of detection (LOD) to 0.3 pM.

Inspired by these results, a second technique was developed by introducing DNA mismatches into the DNA substrate, which offers two advantages compared to using TWJ: (1) the structure of substrate is simpler because mismatch-aided catalytic reaction uses a simple double-stranded duplex containing a mismatch as its substrate, whereas the substrate of TWJ-mediated catalytic reaction needs three strands, or the output needs to contain a stem-loop structure; (2) the use of mismatch allows for modulation of the stability of the substrate. I estimated $\Delta G^o_{f_{Substrate}}$, $\Delta G^o_{f_{Intermediate}}$, and $\Delta G^o_{f_{Product}}$ values and designed substrates with different $\Delta G^o_{f_{Substrate}}$, and studied the relationship between $\Delta G^o_{f_{Substrate}}$ and both the catalyzed reaction and uncatalyzed

reaction, representative of amplification efficiency and background, respectively. The substrate had the largest catalyzed to uncatalyzed rate ratio when $\Delta G^o_{f_{Substrate}} \approx \Delta G^o_{f_{Intermediate}} > \Delta G^o_{f_{Product}}$. The mismatch-aided DNA catalytic reaction achieved a limit of detection of 2.3 pM and a linear range from 5 pM to 500 pM of a short nucleic acid strand.

A third technique using modified nucleotides instead of DNA mismatches was then developed. Three types of nucleotide modifications, including abasic site, anuleotide site, and phosphorothioate linkages were compared. Replacing a nucleotide with an anucleotide site resulted in both high amplification efficiency and low background. This modified nucleotide-aided DNA catalytic reaction successfully demonstrated the detection of microRNA 10b both in buffer solutions and in MB-231 cell lysate samples.

The DNA catalytic reactions reported here opens avenue for future developments, such as intracellular detection, point-of-care testing, DNA computation, and dynamic molecular circuits.

Preface

This thesis is the original work of JingYang Xu. Parts of chapter 1 is published as: Peng, H., Newbigging, A.M., Reid, M.S., Uppal, J.S., Xu, J., Zhang, H., et al. (2019) Signal amplification in living cells: A review of microRNA detection and imaging. Anal Chem, 92, 292. I wrote the section regarding entropy driven catalysis and toehold exchange based reaction and helped revise the manuscript.

Chapters 2, 3, and 4 have yet to be published.

Majority of appendix chapter has been accepted for publication in Biosensor and Bioelectronics. Ethics approval was obtained from the University of Alberta's Research Ethics Board. Xiao, H. & Xu, J.Y., Liu, Y., Feng, W., Pang, B., Tao, J., Zhang, H. (2024). Integration of a Cas12a-Mediated DNAzyme Actuator with Efficient RNA Extraction for Ultrasensitive Colorimetric Detection of Viral RNA. Biosens Bioelectron. 260, 116429. Both Huyan Xiao and JingYang Xu contributed equally to the writing and revision of the manuscript. I designed and conducted the experiments for the extraction concentration and preservation of viral RNA from biological sample. I collected samples from volunteer and analyzed them. I wrote and revised the manuscript and composed the figures. The data presented are entirely my own work.

I have also contributed to writing, composition of figure, sample collection, analysis, and/or method validation for the following publications during my PhD study, which will not be included with this thesis:

1. Liu, Y., Kumblathan, T., Tao, J., **Xu, J.**, Feng, W., Xiao, H., Hu, J., et al. (2023). Recent advances in RNA sample preparation techniques for the detection of SARS-CoV-2 in saliva and gargle. Trends in Anal Chem, 165, 117107.
2. Xiao, H., Hu, J., Huang, C., Feng, W., Liu, Y., Kumblathan, T., Tao, J., **Xu, J.**, et al. (2023) CRISPR techniques and potential for the detection and discrimination of SARS-CoV-2 variants of concern. Trends in Anal Chem, 160, 117000.
3. Deng, W., **Xu, J.Y.**, Peng, H., Huang, C.Z., Le, X.C. & Zhang, H. (2022) DNAzyme motor systems and logic gates facilitated by toehold exchange translators. Biosens Bioelectron, 217, 114704.

4. Liu, Y., Kumblathan, T., Feng, W., Pang, B., Tao, J., **Xu, J.**, Xiao, H., et al. (2022). On-Site Viral Inactivation and RNA Preservation of Gargle and Saliva Samples Combined with Direct Analysis of SARS-CoV-2 RNA on Magnetic Beads. ACS Meas Sci Au, 2, 224.
5. Feng, W., Peng, H., **Xu, J.**, Liu, Y., Pabbaraju, K., Tipples, G., Joyce, M.A., et al. (2021) Integrating Reverse Transcription Recombinase Polymerase Amplification with CRISPR Technology for the One-Tube Assay of RNA. Anal. Chem, 93, 12808.
6. Pang, B., **Xu, J.**, Liu, Y., Peng, H., Feng, W., Cao, Y., et al. (2020). Isothermal amplification and ambient visualization in a single tube for the detection of SARS-CoV-2 using loop-mediated amplification and CRISPR technology. Anal Chem, 92, 16204.
7. **Xu, J.Y.**, Le, C., Chen, B., Zhang, H. (2019). Urban recreational water - potential breeding ground for antibiotic resistant bacteria? J Environ Sci, 81, 1.
8. **Xu, J.Y.**, Zhang, H. (2018). Long-Term Effects of Silver Nanoparticle on the Abundance and Activity of Soil Microbiome. J Environ Sci, 69, 3.

Acknowledgments

I would like to express my sincere gratitude to everyone that has helped and guided me in this journey of learning and discovery. My doctoral study has not only led me to explore and seek new knowledge but also deep into my inner passion.

My deepest gratitude and apologies to my supervisor Dr. Hongquan Zhang for his guidance and patience when I wander astray in my quest for knowledge and independence. Time and time again, His knowledge, expertise, kindness and boundless patience shine through in our discussions when I am troubled, in his encouragement when I am disheartened, and in the completion of this thesis. His mentorship, and exemplary scholarship and citizenship, will continue to guide me in my future pursuit. I wish to extend the same gratitude to my co-supervisor Dr. X. Chris Le for his support and guidance in the completion of this work. His vision and enthusiasm for research help me to see beyond the work in front of me and motivates me to reach for the future. My appreciation and thanks to my current and former supervisory committee members Dr. Andrei Drabovichi and Dr. Xing-Fang Li for their insightful feedbacks, eye-opening perspective and helpful advice and support. The patience and support from Ms. Katerina Carastathis are greatly appreciated.

I am grateful for the help and support of all the postdoctoral fellow and fellow graduate students in the AET division. The stimulating conversation and fun environment allowed me to develop and grow. I especially thank Dr. Jeffery Tao, Dr. Yanwen Lin, Dr. Wei Feng, Dr. Michael Reid, Dr. Ashley Newbigging, Dr. Jinjun Wu, Dr. Bo Pang, Dr. Guan Huang, Dr. Junlang Qiu, Ms. Huyan Xiao, and Mr. Huidong Wang for their tremendous help and friendship.

I would like to thank my church family at Starlight Mandarin ministry at ERBC, for their emotional and spiritual support. I especially thank Pastor Charlene Dong and her family, Her sister

Mrs. Yanling Dong and her family, Mr. Changhao Fang, Mr. Sicheng Quan, Dr. Jie Sun, Mrs Lingzhi Ha and Mr. Jian Se for their friendship, guidance and support.

Most importantly, I would like to thank my parents for their support and motivation. For they help to shape me into who I am, gave me the opportunity to strive for a higher goal, and provided me with a safe harbor to rest and revitalize. Through them I learned it matters not how many times you fall but how many times you get up, and when there is a will there will always be a way. Without their support, I could not have accomplished the work I did.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A	adenine
ABi	abasic site insertion
ABr	abasic site nucleotide replacement
ANi	anucleotide site insertion
ANr	anucleotide site nucleotide replacement
C	cytosine
CHA	catalytic hairpin assembly
Ct	cycle threshold
DNA	deoxyribonucleic acid
dsDNA	double stranded DNA
ssDNA	single stranded DNA
DTT	dithiothreitol
E	amplification efficiency
EDC	entropy drive catalysis
EDTA	ethylenediaminetetraacetic acid
FAM	carboxyfluorescein amidite
G	guanine
ΔG	Gibbs free energy
$\Delta G^o_{f_x}$	free energy of formation of x
ΔG^o_x	free energy change of reaction x
H-bond	hydrogen bond
HCR	hybridisation chain reaction
k	kinetic rate constant
K	equilibrium constant
LOD	limit of detection
MBs	magnetic beads
miRNA	microRNA

nt	nucleotides
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PS	phosphorothioate linkage
ROX	ROX reference dye
RT-qPCR	revers transcription-quantitative polymerase chain reaction
T	thymine
T_m	melting temperature
TMSD	toehold mediated stand displacement
TE	toehold exchange
TWJ	three-way junction

1. Chapter 1: Introduction

1.1 Introduction of DNA

1.1.1 Structure and features of DNA as building blocks for DNA nanotechnology.

Deoxyribonucleic acid (DNA) is a polymer consisting of a backbone of pentose sugars and phosphates, with a purine or pyrimidine base linked to each sugar. DNA naturally exist as a right-handed double helix conformation, with the two strands of DNA running antiparallel to each other. This double helix structure resembles a spiral staircase, with the bases facing each other and pointing towards the center to form the steps, while the sugar and phosphate backbone faces outward, like the railings of a ladder. The purine and pyrimidine bases: adenine (A), guanine (G), cytosine (C), and thymine (T), are paired with each other according to the defined Watson-Crick base pairing rules, where A always pairs with T, and G with C[1-3].

In addition to acting as genetic materials, key characteristics of DNA, such as its stability and the predictability of Watson-Crick base pairing, have provoked the usage of DNA as building blocks for creating nanostructures, nanomachines, and molecular circuits, leading to the field of DNA nanotechnology[5,6]. DNA is incredibly stable as the carrier for hereditary information, demonstrated by the successful cultivation of seeds stored for millennia and inspiring works of science fiction such as Jurassic Park[7]. Similarly, this stability of DNA is crucial for DNA nanotechnology to ensure the robust operation and longevity of the designed DNA nanostructures and nanomachines. Secondly, the predictability of the A-T and G-C base pairing of nucleotides means that through careful design of the DNA sequence, we can reliably control the folding of

DNA into desired 2 and 3 dimensional structures[6,8-10]. These structures can form within a single DNA strand or between multiple strands of DNA. In addition to structural DNA nanotechnology, DNA is also used to construct dynamic DNA devices and nanomachines with designed dynamic functionalities, such as computation and mechanical motion[5].

Advances in vitro DNA synthesis have contributed significantly to the progression of DNA nanotechnology. DNA is synthesized in vitro from 3' to 5', one nucleotide at a time, through multiple cycles of deprotection, washing, nucleotide addition, and polymerization reactions [11,12]. The cost of this operation dropped significantly in the early 2000s to <$1 per nucleotide, making the use of DNA as a nanomaterial practical[6,11,12].

1.1.2 Thermodynamics of DNA

DNA nanotechnology not only depends on the intrinsic properties of DNA, but also builds on the thermodynamics of the interactions between the two strands of the double helix. To construct more complex structures or dynamic nanomachines where strands constantly hybridize and dissociate, we need to understand the forces facilitating the hybridization of the DNA strands to form the double helix. First, hydrogen bonding (H-bond) exists within the AT and GC base pairs, where the AT pair forms 2 H-bonds and the GC pair forms 3 H-bonds. Therefore, the GC base pair contributes more to the stability of the double helix structure. While water can also form hydrogen-bonds with polar groups on the bases, the hydrophobicity of the large planar purine and pyrimidine rings causes internalization of the bases within the double helix structure and form base pairs. This internalization of the bases exposes the sugar-phosphate backbone. As the phosphates are negatively charged at pH 7, the double helix DNA structure is also stabilized by monovalent or divalent cations within the solution serving as counterions. Finally, the double helix is also

stabilized by the Van der Waals interaction between adjacent base pairs due to the stacking of the non-polar planar bases. Although the base stacking interaction between just two adjacent base pairs is weak, it quickly adds up when there are multiple base pairs within the double helix DNA duplex[13].

Specifically, consider thermodynamics of DNA hybridization as a two-phase system (Eq. 1) where two single-stranded DNA (ssDNA), X and its complementary strand X*, hybridize to form a double-stranded DNA (dsDNA), X:X*:

$$X + X^* \leftrightarrow X{:}X^* \qquad (1)$$

Assuming there is no secondary structure associated with the ssDNA X and X*, the standard Gibbs free energy ΔG_f^o is zero for these ssDNA strands. Therefore, the free energy change of the system is equal to the ΔG_f^o of dsDNA, X:X*. The free energy change of the system can thus be described using the following equation:

$$\Delta G^0 = -RT\ln K \qquad (2)$$

Where R is the ideal gas constant, T is temperature in Kelvin, and K is equilibrium constant. K can be described in the following equation:

$$K = \frac{[X{:}X*]}{[X][X*]} = \frac{k_{hyb}}{k_{dis}} \qquad (3)$$

Where [X], [X*], and [X:X*] are the concentrations X, X* and X:X* at equilibrium. k_{hyb} is the rate constant of the hybridization of the ssDNA into the dsDNA duplex, and k_{dis} is the rate constant of the reverse reaction where the dsDNA duplex dissociates into the two ssDNA strands. Therefore, at the melting temperature, T_m, half of the dsDNA would remain as duplexes while the other half would dissociate into the ssDNA. K becomes into the following equation (Eq. 4):

$$K = [X{:}X^*] / ([X]\,[X^*]) = 2\alpha / ((1 - \alpha)^2 Ct) \qquad (4)$$

Where the *Ct* is the total strand concentration within the system, and α is the molar fraction of dsDNA. For example, if we have 2 mM of dsDNA, then *Ct* is 4 mM. Using the relationship between the concentration of the dsDNA and ssDNA, one would theoretically be able to determine the ΔG^{o} of any DNA duplex by measuring the UV absorbance at 260 nM while slowly increasing the temperature of the system. The absorbance increases when dsDNA denatures into ssDNA. However, the spontaneous hybridization and dissociation of DNA occurs over a narrow temperature range. If the temperature is increased by more than 5 °C above the T_m, almost all DNA would be in the single stranded state, and vice versa. Therefore, it is difficult to reliably determine the ΔG^{o} at specific temperatures. Thus, previous studies have found another way to determine ΔG^{o} at given temperature and condition by linking the first equation with the following equation for Gibbs free energy:

$$\Delta G = \Delta H - T\Delta S \qquad (5)$$

To obtain the following equation:

$$\Delta H - T\Delta S = -RT\ln K \qquad (6)$$

By moving T to one side and everything on to the other side, we are left with the following equation:

$$T=\Delta H/(\Delta S-R\ln K) \qquad (7)$$

Thus, we can substitute K for our equilibrium equation determined previously and because at T_m, α=1/2:

$$\frac{1}{Tm} = \frac{\Delta S - R\ln\left(\frac{2\alpha}{(1-\alpha)^2 Ct}\right)}{\Delta H} = \frac{\Delta S + R\ln(Ct/4)}{\Delta H} \qquad (8)$$

Rearranging the equation:

$$\frac{1}{Tm} = \ln(Ct/4)\frac{R}{\Delta H} + \frac{\Delta S}{\Delta H} \qquad (9)$$

After determining the T_m of the DNA at different concentrations, we plot $1/T_m$ on the Y-axis and ln(Ct/4) as the X-axis. The slope is then R/ΔH and the Y-intercept is ΔS/ΔH, which we can use to determine ΔH and ΔS. The ΔH and ΔS are then used to determine ΔG⁰ for the formation of dsDNA, ΔG_f^o, at specific temperatures using equation 5. Using these empirical ΔG_f^o for different DNA duplex with different sequences, researchers developed models for the estimation of ΔG_f^o and T_m from sequences, most notably the nearest neighbour model, which accounts for not only the contribution of individual bases but also the effect its neighbour on the overall thermodynamic stability of the DNA duplex[13,14]. Computation tools such as NUPAC[15] were also developed based on these models to make the calculation easier. These thermodynamic parameters are key considerations when the stability of DNA duplex is evaluated in the design and the development of noncovalent DNA catalytic reactions that have been explored in this thesis.

The model system described here represents an ideal condition, but many factors can affect the thermodynamics of DNA. For examples, depending on the DNA sequence, ssDNA can potentially form secondary structures. These secondary structures can impact both k_{hyb} and k_{dis}. Furthermore, the thermodynamic parameters such as ΔG_f^o, as well as the kinetics rate constants such as k_{hyb} and k_{dis}, are influenced by temperature and buffer conditions, including pH and the valence and concentration of counterions. This is especially true for k_{hyb} and k_{dis} which can differ by orders of magnitude[16]. However, most sources would agree, for most oligonucleotide that is several tens of nucleotides long, a k_{hyb} of $3\text{x}10^6$ $M^{-1}s^{-1}$ is expected for most commonly used buffer conditions[17,18]. This information is crucial when considering strands displacement and the maximum rate at which of strand displacement can occur. As the stability of DNA duplex increase with strand length, concentration, and GC content, one would anticipate a decrease in k_{dis} and ΔG_f^o. Specifically, Morrison and Stols[19] studied both 10-mer and 20-mer and determined their k_{dis} at

30.5 °C in a 1 M NaCl/10 mM NaH_2PO_4 buffer and found the k_{dis} are 1.1×10^{-2} s^{-1} and 1×10^{-10} s^{-1}, respectively. This means the dissociation half-life of DNA duplex shorter than 10nt is less than one-minute while a half-life of >100 years can be expected for duplex with similar composition of bases that is longer than 20nt. This finding and the correlation between ΔG_f^o and k_{dis} outline in equations 1 and 2 shown above provide useful guide for designing the oligos and duplex used in DNA nanotechnology.

1.2 Toehold-mediated strand displacement reactions

Toehold-mediated strand displacement (TMSD) relies on a motif known as toehold, which is a short ssDNA region extending from the end of the dsDNA. Toeholds are typically 6-8 nucleotides (nt) long. Like sticky ends facilitating the colocalization of two pieces of dsDNA for DNA ligation, the toehold enables the invading strand, **I**, to colocalize with the dsDNA, **R**, and promote the displacement of the output strand, **O** (**Figure 1.1**). The reactant duplex **R** is comprised of a **O** strand and its complement **C**. The DNA strands consist of different functional domains labeled by numbers. Domain **n** is complementary to domain **n***. The arrow denotes the 3' end of a DNA sequence. The toehold is the single-stranded overhang domain **1*** at the 5' end of **C**. The displacement reaction starts with binding of domain **1** of **I** with the toehold **1***, colocalizing of **I** and **R**, which initates a branch migration to displace **O** by **I**[20]. Based on the k_{hyb} and k_{dis} discussed in previous section and from literature[19, 21, 22], the half-life of the toehold binding is expected to be on the scale of milliseconds. This brief association is sufficient to increase the attempt frequency for initiating displacement through branch migration. The branch migration is a random walk process, during which **I** and **O** compete for the same nucleotides on domain **2*** of **C**. When DNA breathing occurs at the 5' end of domain **2***, leading to breaking of some adjacent H-bonds between

2* and **2**, the base on **C** is then available to form new H-bonds with either **I** or **O**. During the branch migration, the breaking and forming of H-bond is completely random, and each step of this random walk can have 2 outcomes. First, the bond between **O** and **C** breaks, and new bond form between **I** and **C**, and the reaction moves to the right. Second, the bond between **I** and **C** can break and new bond form between **O** and **C**, and the reaction moves to the left. While each step of this random walk takes only 12-20 μs[23-25], the overall branch migration time is much longer due to the randomness of the walking process and depends on the number of steps involved. Branch migration completes when **I** fully hybridizes with **C**, forming a new DNA duplex **P** and displacing **O**. Because there is no reverse toehold for binding of **O**, the rate constant for the reverse reaction, k_r, is orders of magnitude smaller than that of the forward reaction. Therefore, the presence of the toehold drives the reaction towards the formation of **O** and **P**, rendering the toehold-mediated strand displacement reaction irreversible.

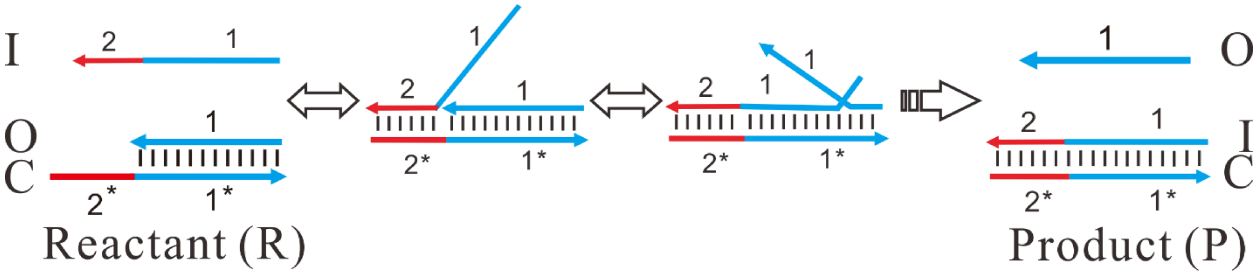


Figure 1.1 *Toehold-mediated strand displacement between a ssDNA and a dsDNA duplexes. The toehold on the reactant duplex is a single-stranded overhang domain 1*. The invading strand, dented as **I**, displaces the output strand, denoted as **O**, to form the product duplex. C demotes the complementary strand.*

The stability of the hybridization between the toehold and its binding region determines the kinetics of TMSD. The duration of the colocalization is determined by the stability of the

toehold. The hybridization and dissociation rate constants k_{hyb} and k_{dis} of the toehold is related to the equilibrium constant K, and the free energy of hybridization ΔG_f^o as outline by equations 2 and 3. Therefore, shorter toeholds mean a higher ΔG_f^o for hybridization of the toehold with its binding domain. The hybridization is less stable and the colocalization half-life is thus shorter, while longer toeholds result in a lower ΔG_f^o for the hybridization, which is more stable and the colocalization half-life is thereby longer. The ΔG_f^o of the toehold is correlated with its length and GC content. As the toehold improves kinetics of TMSD by increasing attempt frequency of the branch migration, one would expect a relationship exists between the ΔG_f^o of the toehold and kinetics of TMSD. Indeed, both Yurke & Mills[18] and Zhang & Winfree[24] studied the relationship between the length of the toehold and the rate constant of the displacement reaction. Strikingly, the rate constant (k_f) of TMSD is increased by 6 orders of magnitude when the toehold length is increased from 0 to 6 nt. Two groups found that there is an almost linear relationship between the base-10 logarithm of k_f (or k_{eff} in **Figure 1.2**) and the length of the toehold when the length of the toehold is shorter than 6 nt. Furthermore, Zhang & Winfree revealed that 6-7 nt was the saturation point for toehold length. Zhang & Winfree also studied the relationship between ΔG_f^o of the toehold and k_f and found there is an almost linear relationship between them and k_f reaches a saturated value when ΔG_f^o is lower than -8 kcal/mole[24].

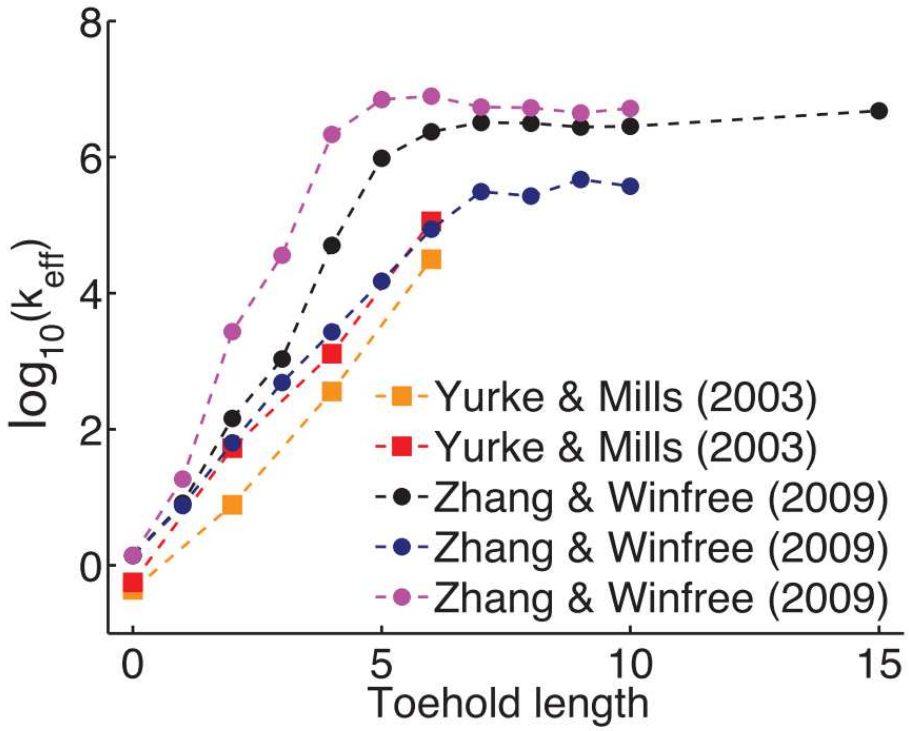


Figure 1.2 *Correlation between the effective rate constant of TMSD k_{eff} (or k_f) and toehold length. Data of Yurke and Mills[18] and the data of Zhang and Winfree[24] are overlay and the log based 10 of k_f is plotted again the toehold length in nucleotides from 2 separate studies. Different toehold and branch migration sequences are used for these studies. For Zhang and Winfree, the samples were run at 25 ºC with 12.5 mM Mg^{2+}, the purple dots are data for a GC rich toehold, the dark blue dots are for an AT rich toehold and the black dots are for when there is a mixture of bases in the toehold. The same branch migration region is used for all three data set. Yurke and Mills[18] performed their experiment at 20 ºC with 1M Na^{+} and did not make distinction between the two toeholds and branch migration region used. Reprinted with permission from ref 25. Copyright 2013 Oxford University Press.*

The extensive simulation and energy landscape modeling conducted by Srinivas et al.[25] help to further explain the kinetic of TMSD, especially the importance of toehold for initiation of the branch migration and TMSD. They found that the time of DNA breathing is significantly shorter than the step rate of random walking. They also suggested the presence of a potential energy barrier that must be overcome to initiate the branch migration. Further simulation using a

coarse-grained DNA model by oxDNA[26-28] suggests the reason why the step rate is slower than DNA breathing lies in the fact that the advancement of the branch point requires substantial structural rearrangement of the two competing strands. This is because the free floating single stranded branches are repelled and bend away from each other. The repulsion between the single stranded overhangs of the two competing strands, **I** and **O**, also contribute to the energy barrier for the initiation of the branch migration because the bent DNA strands cause partial disruption of base stacking and restrict the conformational freedom at the branch point. These findings and the results from Zhang and Winfree[24], all indicates the toehold provides the activation energy needed to overcome the energy barrier and initiates the branch migration and TMSD. The contribution of the toehold to TMSD can thus be explained using the Arrhenius equation:

$$k_f = Ae^{-E_a/RT}$$

Where E_a is the activation energy, estimated by using toehold binding energies[22]. Toehold binding energy can be estimated by many web-based software, such as NUPACK[29]. A is the rate constant when toehold length is zero.

The rate constant for the TMSD reaction can be determined experimentally using the rate law and in-silico modeling. The rate of the forward reaction as function of k_f can be described in the following equation, using the reaction in figure 1.1 as an example:

$$\text{Rate} = -\frac{d[S]}{dt} = -\frac{d[I]}{dt} = \frac{d[P]}{dt} = \frac{d[O]}{dt} = k_f[S][I]$$

If **S** and **I** are present at an equal amount :

$$\frac{d[S]}{dt} = k_f[S]^2$$

$$d[S] = k_f[S]^2 dt$$

after integration of the equation we would be able to find

$$\frac{1}{[S]} = \frac{1}{[S]_0} + k_f t$$

However, if [S]>>[C], the equation describe in the previous section is not applicable and needs to be modified for when S and I are present at different amounts:

$$[S] = ([S]_0 - x); [I] = ([I]_0 - x)$$

$$\frac{d([S]_0 - x)}{dt} = -k_f([S]_0 - x)([I]_0 - x)$$

$$-\frac{dx}{dt} = -k_f([S]_0 - x)([I]_0 - x)$$

$$\frac{dx}{([S]_0 - x)([I]_0 - x)} = k_f dt$$

After integration we end up with the following equation:

$$ln\frac{[I][S]_0}{[S][I]_0} = k_f([I]_0 - [S]_0)t$$

Zhang & Winfree[24] also developed in silico method for determining k_f. They established a series of differential equations to define the change in concentration of each of the reactants and product during the reaction, and provided a range for the estimation of k_f and generated a model for the reaction. The model is able to simulate the consumption of reactant and production of product given an input of the initial amount reactants and product for an estimated k_f. Subsequently, they calculated the absolute difference between each point of the modeled results and the experimental data, thereby determining the total error of the model as the sum of these differences. Finally, they utilized a function to identify the model with the lowest total error, obtaining its corresponding k_f as the result for the best fit k_f value.

A double TMSD can also occur between two dsDNA resulting in an exchange of DNA strands and the formation of two longer DNA duplexes. This process involves the formation and

migration of a Holiday junction or a 4-way junction[30, 31]. As shown in **Figure 1.3**, the two DNA duplexes each have two toeholds. The left duplex has domains **1*** and **3** as its toeholds, while the right one has domains **1** and **3*** as toeholds. The complementary toeholds, **1** and **1***, and **3** and **3***, hybridize to form the Holiday junction shown in the bottom left. The formation of the Holiday junction initiates branch migration and strand displacement. At the junction point, base pairing interactions are constantly broken and reformed between different strands at random. With each step of this random walk, either the arms with the toeholds are lengthened by one base and the other two are shortened, or the opposite could occur. When the branch point reaches the end of the duplex distal to the toehold, the complex dissociates and forms two longer dsDNA products. The hybridization of the two toehold pushes the equilibrium toward the longer product duplexes, which are more energetically stable. Studies have shown the valency of the counterions in buffer has a strong effect on the reaction rate and decrease in temperate drastically reduces the reaction rate[32]. At 37 °C, mean step time is 0.3 s when the buffer contains Mg^{2+} while the same buffer containing just Na^{+} the step time is only 0.3 ms.

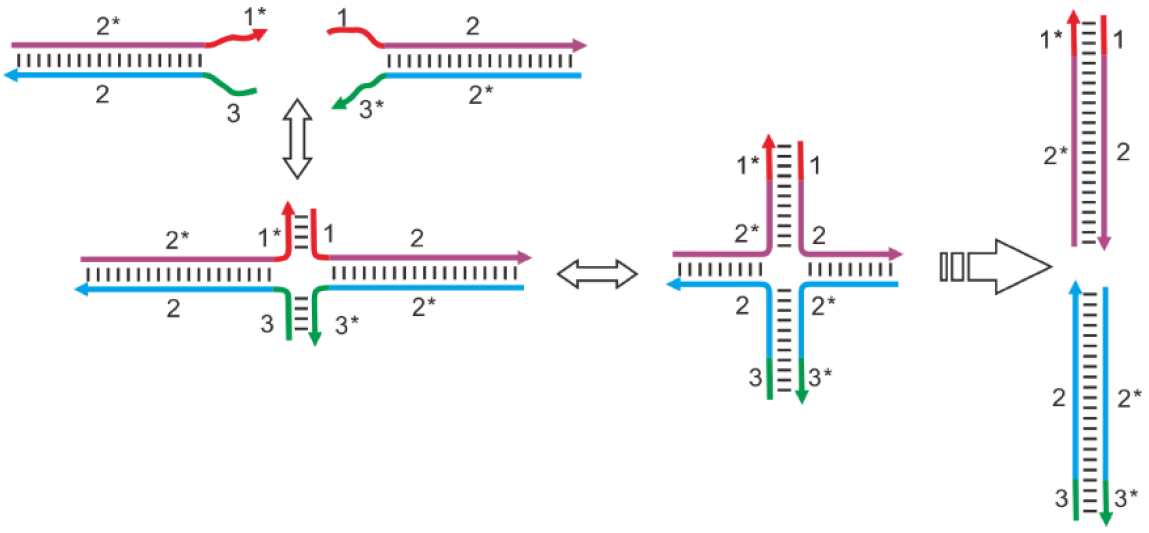


Figure 1.3 *Toehold-mediated strand displacement between two DNA duplexes.*

The two DNA duplex has identical dsDNA region each with two toeholds that is complementary to the toeholds on the opposite DNA duplex. The hybridization of the toeholds to its complementary toehold on the other duplex forms the Holiday junction shown on the lower left. Through random walking the Holiday junction would migrate and eventually result in the two new duplex products with more complementary base pairs on the far right.

1.3 Toehold exchange reaction

The toehold exchange (TE) reaction is a variant of TMSD where a new toehold is formed upon displacement of the output strand at the end opposite to the initial toehold. As shown in **Figure 1.4**, The initial toehold, domain **1*** in **R**, is exchanged into a new toehold, domain **3*** in **P** after TE reaction. Specifically, the invading strand **I** bind to the initial toehold, domain **1*** on **C**, initiating the branch migration to displace **O**. Once **I** completely hybridizes with domains **1*** and **2*** of **D**, **O** either spontaneously dissociates from **C**, or through its interaction with **3*** on **C** facilitates the reverse reaction. If **O** dissociates, **I** and **C** forms the duplex product **P**, with a new toehold, domain **3***. Thus, the initial toehold, domain **1***, in **R** is traded for this new toehold domain

3* in **P**. The displaced **O** can also bind to this toehold to facilitate the reverse reaction. In TE, the reverse reaction can occur at a reasonable rate due to the presence of the reverse toehold.

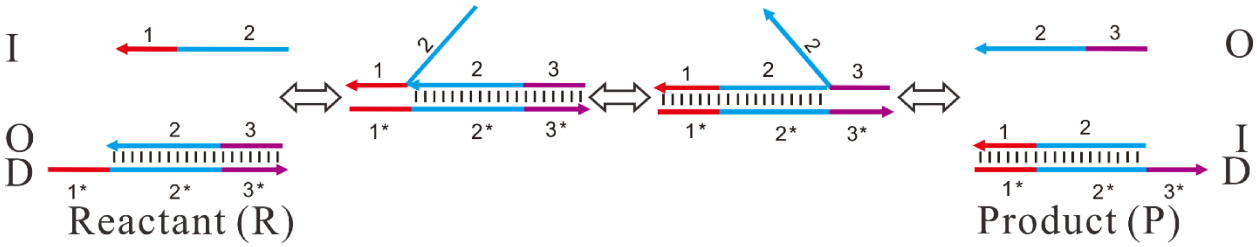


Figure 1.4 *Toehold-mediated strand exchange or toehold exchange reaction.*
*There is an exchange of the toehold **1*** of the reactant, **R**, for the new toehold **3*** of the product, **P**. In the forward reaction, the invader strand, denoted as **I**, invades **R** via the toehold **1*** and displaces the output strand, denoted as **O**, forming **P**. This exposes a new toehold **3***, where **O** can bind to facilitate the reverse reaction. The direction of equilibrium depends on the ΔG^0 and length of **1*** and **3***, as well as the concentrations of **I** and **O**. **D** denotes the toehold-containing strand.*

The equilibrium of the TE reaction is determined by both the concentration of the reactant and the ΔG_f^o between the initial and reverse toeholds. If **1*** and **3*** are identical in sequence and length, the binding energy is the same for two toeholds. Therefore, there is no bias in the direction of the equilibrium and K=1. The equilibrium state is entirely based on the concentration of reactants, where if one is in excess, the equilibrium will shift to the opposite end and consume the reagent in excess until a new equilibrium is reached. For example, if at the beginning of the reaction $[\mathbf{O}]_0$ and $[\mathbf{P}]_0$ are both zero, and $[\mathbf{I}]_0=[\mathbf{R}]_0$, then the system would favour the formation of **O** and **P** and the concentration of each component would eventually reach an equilibrium concentration of $[\mathbf{I}]_0/2$. While if $[\mathbf{I}]_0>>[\mathbf{R}]_0$, then at equilibrium $[\mathbf{I}]_{eq}\approx [\mathbf{I}]_0$ and $[\mathbf{R}]_{eq}\approx 0$, and $[\mathbf{O}]_{eq} \approx [\mathbf{P}]_{eq}\approx[\mathbf{R}]_0$. The high concentration of **I** is thus able force the equilibrium towards the formation of **P** and **O**. However, if one toehold is longer than the other or has a lower binding energy due to

the sequence of the toehold such as a higher GC content, the equilibrium would favor the formation of the duplex with the shorter toehold. For example, if **1*** has a lower ΔG_f^o than **3***, k_f would be bigger than k_r and K>1, and vice versa.

The relationship between the length of the toehold and kinetics of the TE reaction was also explored by Zhang & Winfree[24] (**Figure1.5**). The results showed that the length of the forward toehold, **n** in **Figure 1.5**, is the main determinant of k_f. When the length of the forward toehold is equal to longer than that of the reverse toehold, the k_f is mostly determined by the length of the forward toehold, and the k_f of the TE reaction is comparable to those of TMSD reaction. However, when the reverse toehold is longer than the forward toehold, the k_f decreases exponentially[27].

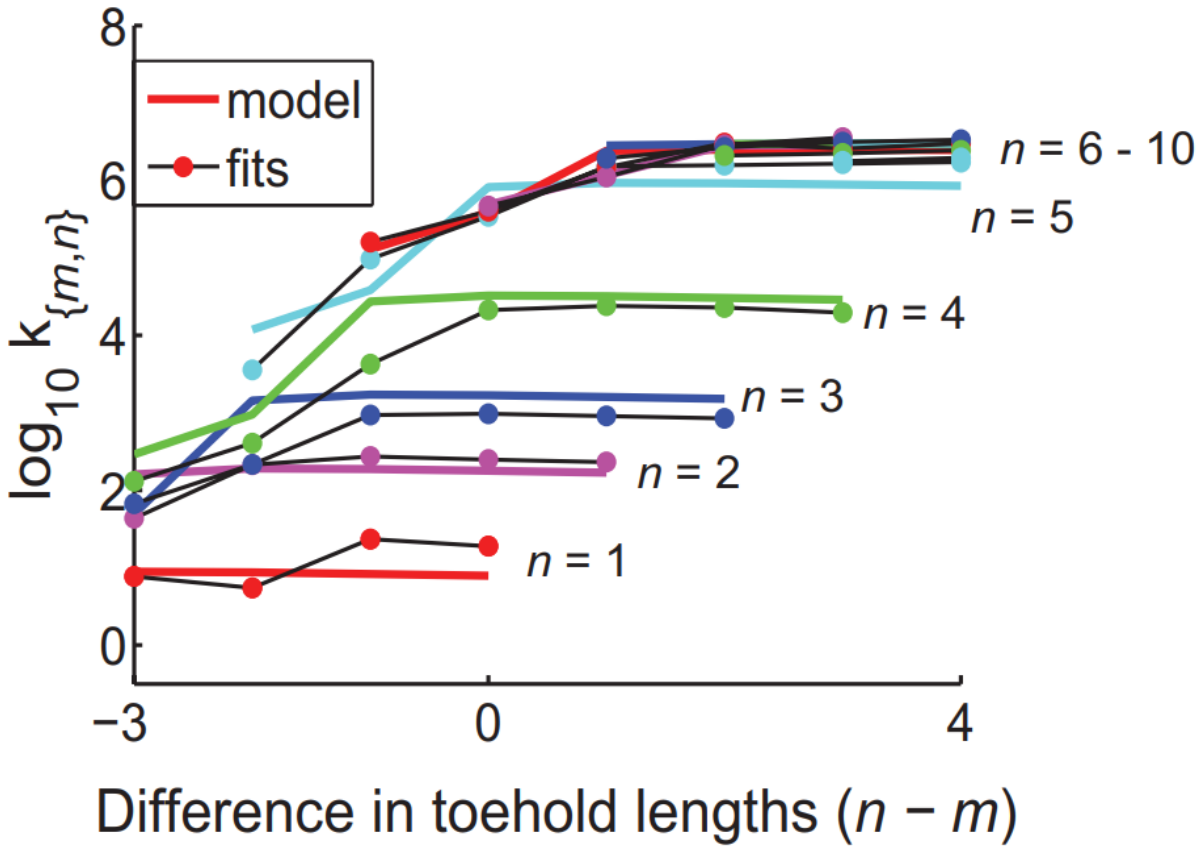


Figure 1.5 *Rate of toehold exchange as a function of the difference in toehold length. Here, n and m denote the length of the forward and reverse toeholds in the number of nucleotides, respectively. The solid lines labeled "model" is a prediction for k based on an earlier model from TMSD. The dots labeled "fits" are best-fit rate constants determined experimentally. Reprinted with permission from ref 24. Copyright ©2009 American Chemical Society.*

Experimental determination of k_f of TE should consider the fact that TE is reversible. First, to favor the forward reaction and push the equilibrium to the right or product side, one reagent is often added in excess of the other reagent. Second, to prevent the reverse reaction and continuously promote and monitor the forward reaction, the products **O** or **P** are consumed in a second irreversible reaction. For example, the output can participate in the signal generation reaction and react with a TMSD reporter. The in silico modeling developed by Zhang & Winfree[24] can utilize this method, where the k_f of the signal generation reaction, k_{rep}, is first determined. The k_{rep} is then

inputted into the algorithm as a constant to determine k_f of the TE reaction. Otherwise, the determination of k_f for the TE reaction is the same as those for the TMSD reaction.

The TE reaction greatly expands the versatility of the TMSD reaction and enables the development of many non-covalent DNA catalytic reactions[5,6]. Compared to the irreversible TMSD where the lack of the reverse toehold prevents the regeneration of invading strand, TE is able to regenerate the invading strand for constructing non-covalent DNA catalytic reactions. Another important feature of TE is that the reverse reaction can be facilitated by other DNA strands besides the output strand. The reversibility and programmability of TE, and the directionality and irreversible nature of TMSD are all important qualities when designing non-covalent DNA catalytic reactions, building molecular logic gates, realizing molecular computation, and detecting disease biomarkers.

1.4 Dynamic DNA nanotechnology

DNA nanotechnology is a diverse and emerging field that utilizes DNA as building blocks to create constructs in the nanometer scale[5, 6]. The unique properties of DNA as a nanomaterial, outlined in section 1.1, have only been exploited in recent years. This is because the advances in automated in vitro DNA synthesis, which also simplifies the process of incorporating fluorescence labeling and chemically modified nucleotides into DNA oligos, has made the use of DNA oligonucleotide economical. Thus, long chains of DNA oligos with desired sequences and modifications can now be synthesised at a fraction of the previous cost. Dynamic DNA nanotechnologies focus on construction of DNA molecules with moving parts and reaction networks that operate on the hybridization between different DNA motifs. In particular, the focus

of this thesis is on non-covalent reaction networks and amplification cascades[5, 6, 33] that operate based on TSMD and TE reactions.

1.4.1 DNA Tweezers

In 2000, Yurke et al.[34] demonstrated the construction of a first dynamic DNA device termed DNA tweezer whose conformation can be changed and subsequently restored using TMSD. This DNA tweezer can switch between its closed and open states via the introduction of a fuel strand and the removal of fuel through hybridization with a complementary strand via TMSD. The design and operation of this DNA tweezer is shown in **Figure 1.6**. This early attempt of imitating biological proteins with DNA polymers opens avenues for construction of other artificial molecular devices.

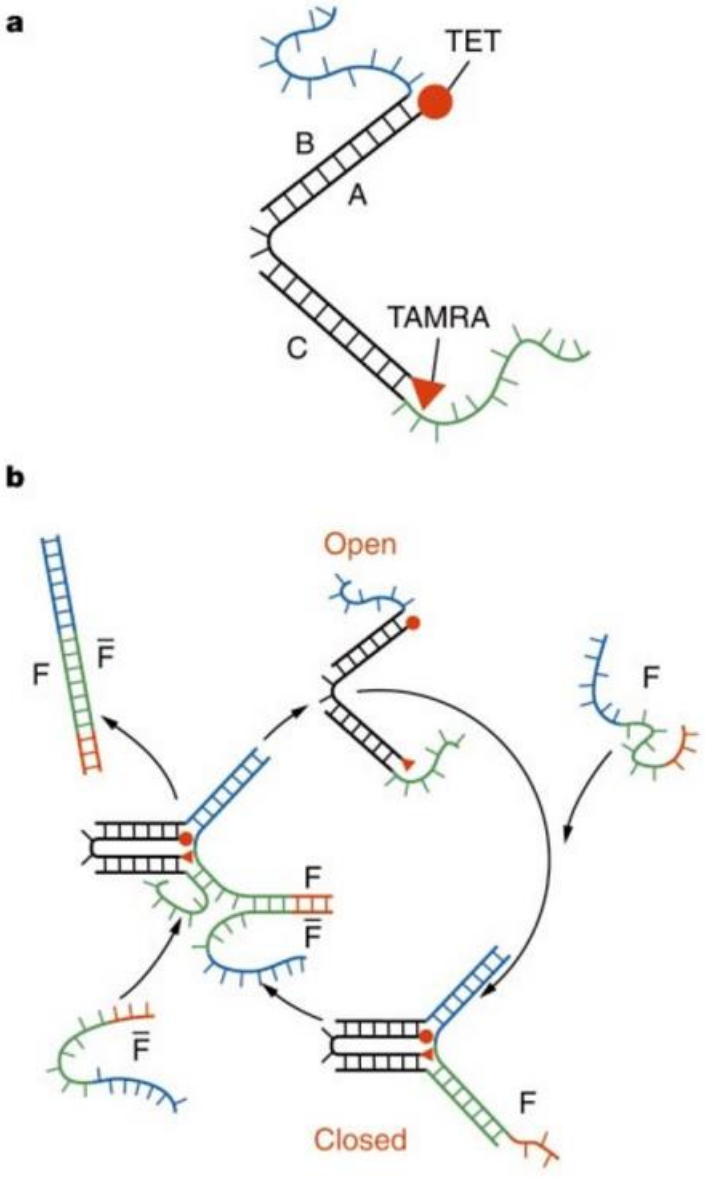


Figure 1.6 *Schematics and design of the DNA tweezer.*

(A) the DNA tweezer is a hybrid of the strands A, B and C. The strand B and C binds the ends of A to form two double stranded arms, while the medial region of A is unhybridized, forming the flexible single stranded hinge. (B) The operation of the DNA tweezer. The tweezer is closed when a fuel strand F binds to the free single stranded overhang of B and C (shown in blue and green) and pulling the two arms together. The $\bar{F}$ *restores the tweezer to it open state by hybridizing with the toehold of F (red) and removing it from the tweezer through TMSD, forming a waste product* $F\bar{F}$*. Reprinted with permission from ref 34. Copyright ©2000 Springer Nature.*

1.4.2 DNA motors

The motion or movement of dynamic DNA systems can be induced through varies means. Protein motors, such as kinesin, myosin, and dynein motors, can walk autonomously along tracks microtubules and actin[35, 36]. The free energy resulting from hydrolysis of ATP into ADP is used to power the walking of protein motors. Inspired by protein motors, researchers constructed a variety of synthetic DNA motors to mimic walking of proteins motors. While initially constructed DNA walkers successfully demonstrated the stepwise and directional walking of a DNA motor along a DNA track, they were unable to achieve the same autonomy of their protein counterparts[35, 36]. An example of this DNA motor reported by Shin and Pierce[36] used TMSD to control bi-pedal motion along a DNA track (**Figure 1.7**.) The success of these early DNA motors sparked interest in creating other DNA motors that can autonomously walk along a DNA track, powered by enzymes, including nicking enzymes and DNAzymes.

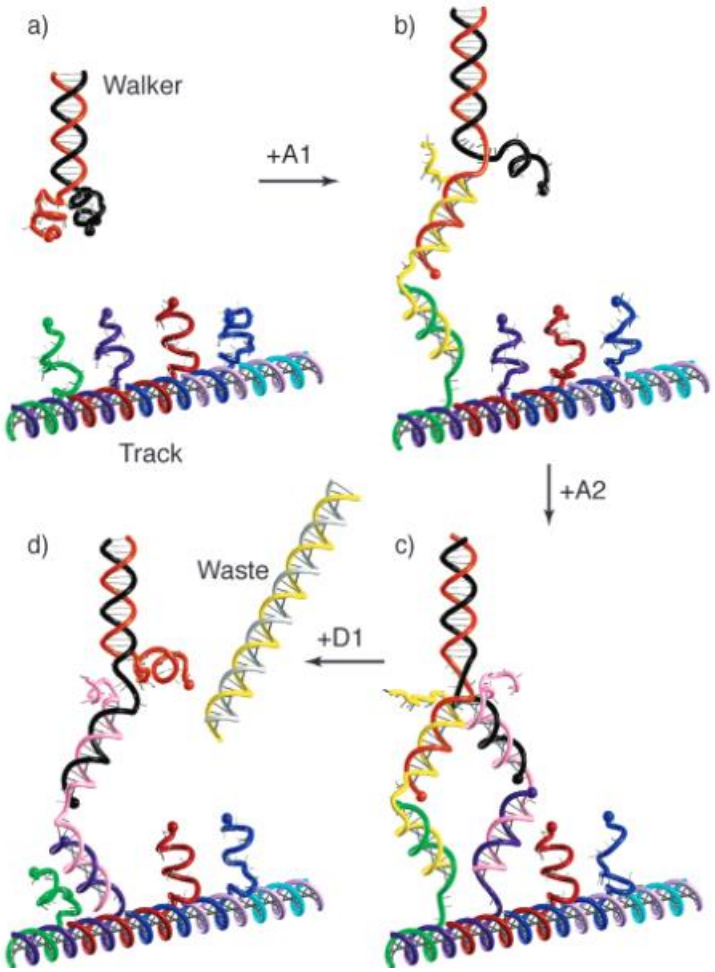


Figure 1.7 *Operation of DNA walker through TMSD.*

(a) The unbound walker and track. Through the addition of a linker strand A1, (b) the walker (first "leg", orange) is attached to the first branch on the track (green). Next, the linker A2 is added, and (c) the second "leg" of the walker(black) to be attached to branch 2 (purple) on the track (purple). Finally, D1 is added to the reaction, and (d) the duplex waste (A1:D1) is produced through TMSD and first "leg" of the walker is released from branch 1. The released leg of the walker can then make another step through the addition of more linker strand. It can either make a step forward and attach to branch 3 through a A3 linker or backward through addition of more A1 linker. The branches on the tracks and legs of the walkers are color code for the dyes (shown as spheres, green=HEX, purple=CY5, red=FAM, blue= Texas Red) and quencher (also spheres, orange=BHQ1, and black= IBRQ) attached to it for detection of walker movement. Reprinted with permission from ref 36. Copyright ©2004 American Chemical Society.

1.4.3 DNA circuits

DNA circuits are constructed for realizing DNA computation and processing complex information. A variety of decision-making tools such as logic gates have been designed and constructed to release an output DNA upon an input DNA via TMSD and TE[37-39]. Information can be stored in the concentrations, spatial localizations and/or chemical properties of molecules; processing of theses molecular information is mediated by chemical reactions as well as TMSD or TE reactions[37, 38]. DNA circuits have been designed to perform computations such as playing a tic-tac-toe game. **Figure 1.8** shows an example of DNA circuits constructed for achieving a AND logic gate[39]. The details of DNA circuit are not discussed here as other reviews [37, 38] already provide lots of insightful discussion for these systems.

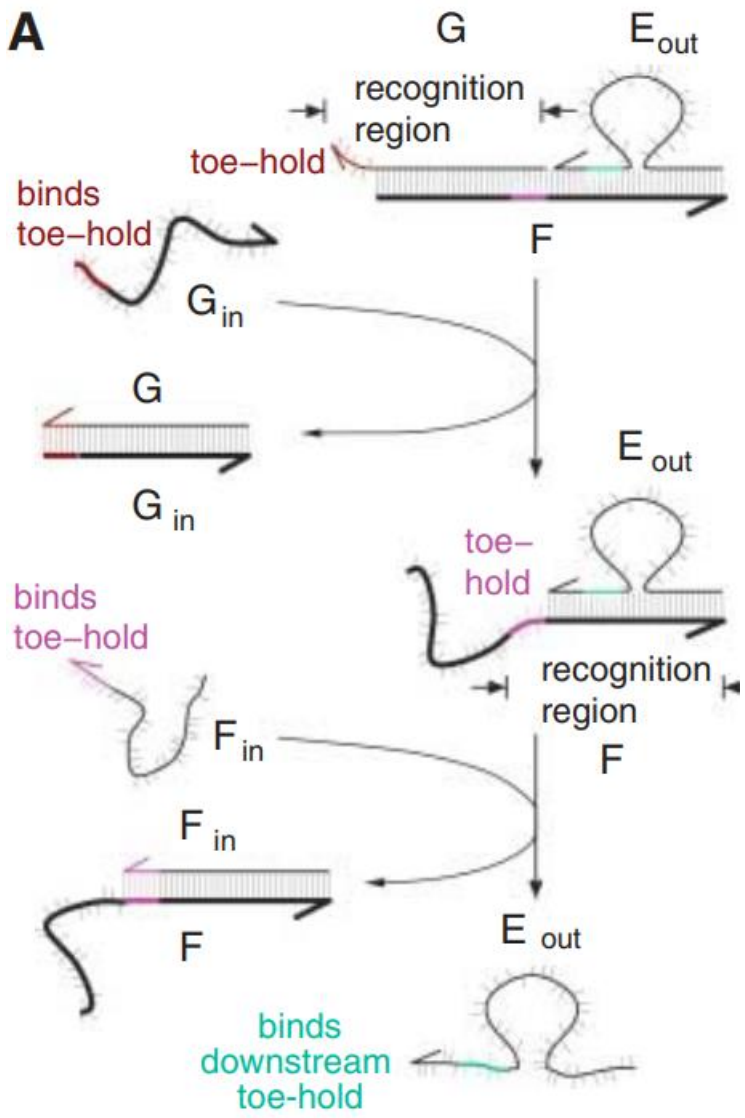


Figure 1.8 *Example of an AND logic gate.*

For AND logic gate only when both inputs, Gin and Fin, are present will the output, Eout, be released to enable further reaction. Reprint with permission from ref 39. Copyright © 2006, The American Association for the Advancement of Science.

1.4.4 Noncovalent DNA catalytic reactions

The breakdown of ATP by intrinsic ATPase activity fuels molecular motors, which inspired design of enzyme-free, non-covalent DNA catalytic reactions to power walking of DNA motors[40]. Unintentionally, the attractive capability of enzyme-free, non-covalent DNA catalytic reactions to facilitate chemical and signal amplification has led to their exploration for biosensing applications, which has become the focal point of further development.

Initial attempts to imitate catalytic breakdown of ATP began by designing ATP-like surrogate molecules in the form of metastable DNA complexes. These complexes can transit to a lower energy configuration in the presence of a catalytic DNA, where the catalytic DNA initiates the catalytic reaction through reacting with a complex molecule. The conformation of the complex then collapses into a final product, releasing the catalytic DNA[40, 41]. The released catalytic DNA then reacts with another complex molecule, leading to another round of reaction. This process repeats, achieving continuous catalytic reactions and signal amplification. To date, a variety of noncovalent DNA catalytic reactions have been developed. We categorize them into three groups based on how TMSD and TE reactions are involved.

1.4.4.1 Noncovalent DNA catalytic reactions based on TMSD.

Controlling access to the toehold is an important strategy in the design of non-covalent DNA catalytic reactions. For example, Dirk and Pierce developed a hybridization chain reaction (HCR) by using a DNA hairpin with a long stem and a short loop to kinetically trap the toehold within the loop of DNA hairpin[42].

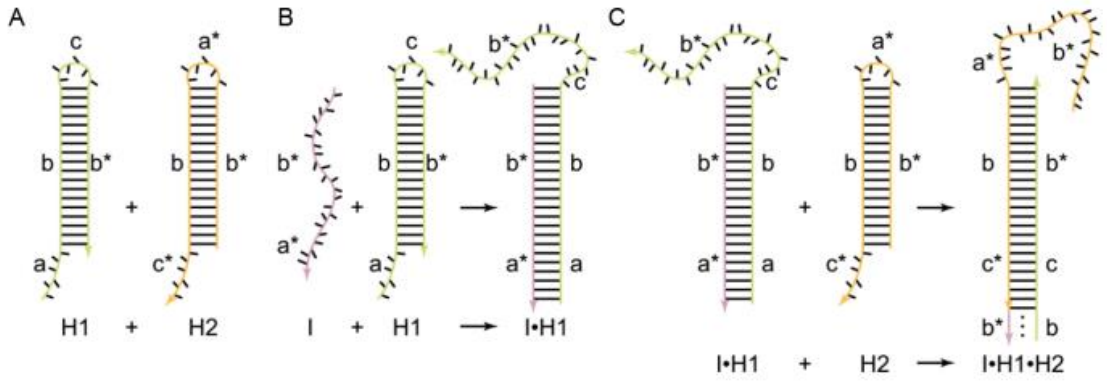


Figure 1.9 *Schematics of hybridization chain reaction (HCR).*

(A) A metastable state exist substrates hairpins H1 and H2, as the long stems kinetic traps the toehold in the loop region to restrict its access. (B) Initiator strand I bind the toehold a on H1 to initiate toehold-mediate strand invasion and open H1. (C) The freed single-stranded overhang of the I·H1 hybridize with hairpin H2 by invasion via toehold c. This polymerization process continues through an alternating step-wise addition of one H1 followed by a H2 hairpins. For sensor applications, H1 and H2 can be designed such that the analyte can function as initiator, or a translator that releases the initiator in the presence of the analyte can be used. Reprinted with permission from ref 42. Copyright 2004 National Academy of Sciences.

In HCR, the substrates are two stable hairpins, **H1** and **H2**, both with a long stem and a short loop region (**Figure 1.9**). The short overhangs at the 5' end of **H1** and 3' end of **H2** are the toeholds. The toehold in one hairpin is complementary to the loop region of the other hairpin. The stem regions are identical for **H1** and **H2**. The initiator strand, which is complementary to the toehold and stem of the hairpin, triggers HCR by opening one of the hairpins through toehold mediated strand displacement. Although the reaction of **H1** with **H2** is energy favorable, this reaction is very slow because of their hairpin structures[42]. Upon the addition of an initiator strand, it reacts with the stem of **H1** through TMSD, opening the hairpin and exposing a single-stranded sequence. This sequence then acts as the initiator to react with **H2** through another TMSD. Similarly, the initiator sequence for **H1** is then exposed from **H2**. This process continues, leading

to a chain reaction where **H1** and **H2** are assembled in an alternating approach to form a long double helix concatemer.

The hybridization between the toehold and loop region contributes to the thermodynamic driving force primarily through enthalpic effects, facilitated by additional base pairing and base stacking interactions. Although opening the hairpin loop can contribute to entropic gains, such gains are arguably lost as the formation of the large concatemer decreases the moieties in the reaction system. Thus, an overall loss in entropy could partially explain the slow rate of reaction and poor amplification efficiency for HCR. Although HCR often afford only 6-10 fold signal amplification in one hour, it is advantageous for intracellular and single molecule detection[33] as the colocalization of initiator and concatemer enables the localization of signals around the target. Ideally, the targets are single-stranded nucleic acids with very little or no secondary structures. This enables efficient hybridization with the substrate to initiate the reaction. Common targets are messenger RNA or miRNA[33].

A four-way branch migration variant of HCR[44] has also been reported. In this design, a dsDNA duplex AR initiates the reaction between **H1** and **H2** through formation of a mobile Holliday junction with **H1 (Figure 1.10)**. This is facilitated by sticky ends **a*** and **x** of **AR** that bind to their toeholds **a** and **x*** on **H1** (state 1 in **Figure 1.10**). The complex transforms into state 2 through four-way branch migration, opening the hairpin **H1** and exposing the loop domain **c**. Domains **c** and **y** can now bind their toeholds on **H2** to form another Holliday junction intermediate (state 3). Progress of state 3 into state 4 liberates two new domain **x** and **a***, which can then bind to their toeholds in another **H1**. This sequential hybridization of **H1** and **H2** repeats cyclically to generate a long concatemer. However, compared to TMSD, the branch migration of the Holliday junction has a much slower step rate, and the overall reaction rate is slow. This reaction has been

used as a sensor for detection of dsDNA. It was also used to simulate actin polymerization motor of Rickettsia bacteria[45] and perpetuate the swelling of DNA-cross-linked hydrogels[45, 46].

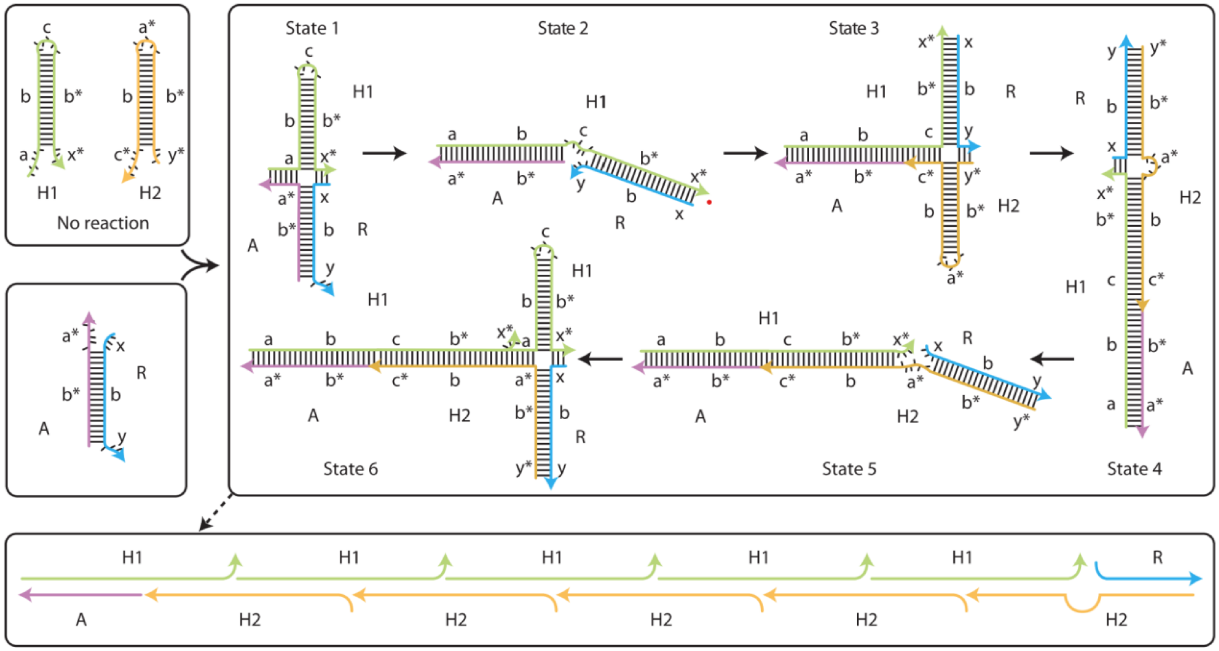


Figure 1.10 *Variant of hybridization chain reaction employing four-way branch migration.*
The initiator complex forms a 4-way branched junction with ***H1*** *for state 1. Migration of the Holliday junction produces the intermediate shown in state 2.* ***H2*** *binds with the intermediate complex to form state another 4-way branched junction in state 3, which transforms into state 4 and 5 similar to how state 1 transforms into state 2. The cycle repeats when another* ***H1*** *is recruited onto step 5 to form state 6 and the cyclic sequential addition of* ***H1*** *and* ***H2*** *in a polymerization chain reaction. This reaction is achieved through the constant toehold exchange reaction with the strand* ***R****, where either the x or y domain toehold is exposed along with either domains* ***a**** *or* ***c*** *to enable reaction either with* ***H1*** *or* ***H2*** *respectively. Reprinted with permission from ref 44. Copyright 2007 Springer Nature Limited.*

To improve amplification efficiency while maintaining the highly concentrated signal, researchers developed a branched HCR (**Figure 1.11**)[47, 48]. Branching and dendritic growth of HCR products is achieved using specially designed substrates that creates 2 identical toeholds during each cycle. Instead of DNA hairpins, the substrates used are dsDNA with either 1 or 2 loops

in the middle of the duplex. The strands are labelled with a pair of fluorophore and quencher for signal generation. The trigger input initiates the reaction through its binding to the double-looped substrate-A, opening the first loop through TMSD. This exposes a previously blocked toehold in the quencher-labeled strand, where a fuel strand Assistant-A can bind and hybridize with the quencher-labeled strand through TMSD, forming Byproduct-A. Thus, two single-stranded loops are then liberated for facilitating binding with two substrate-B molecules. Such binding also released a previously blocked toehold from each substate-B molecule for the second fuel strand Assistant-B to hybridize with the quencher-labeled strand in substrate-B though TMSD, resulting in a dsDNA product with two single-stranded branches. These single-stranded branches then react with two substrate-A molecules, leading to the formation of two dendritic branches from each old branch. Therefore, exponential amplification can be achieved, forming a dendrimer product. This exponential amplification enables the branched PCR to have a much stronger amplification power than the conventional HCR. However, the use of duplex substrate with two single-stranded loops makes the reaction more complicated and prone to leakage.

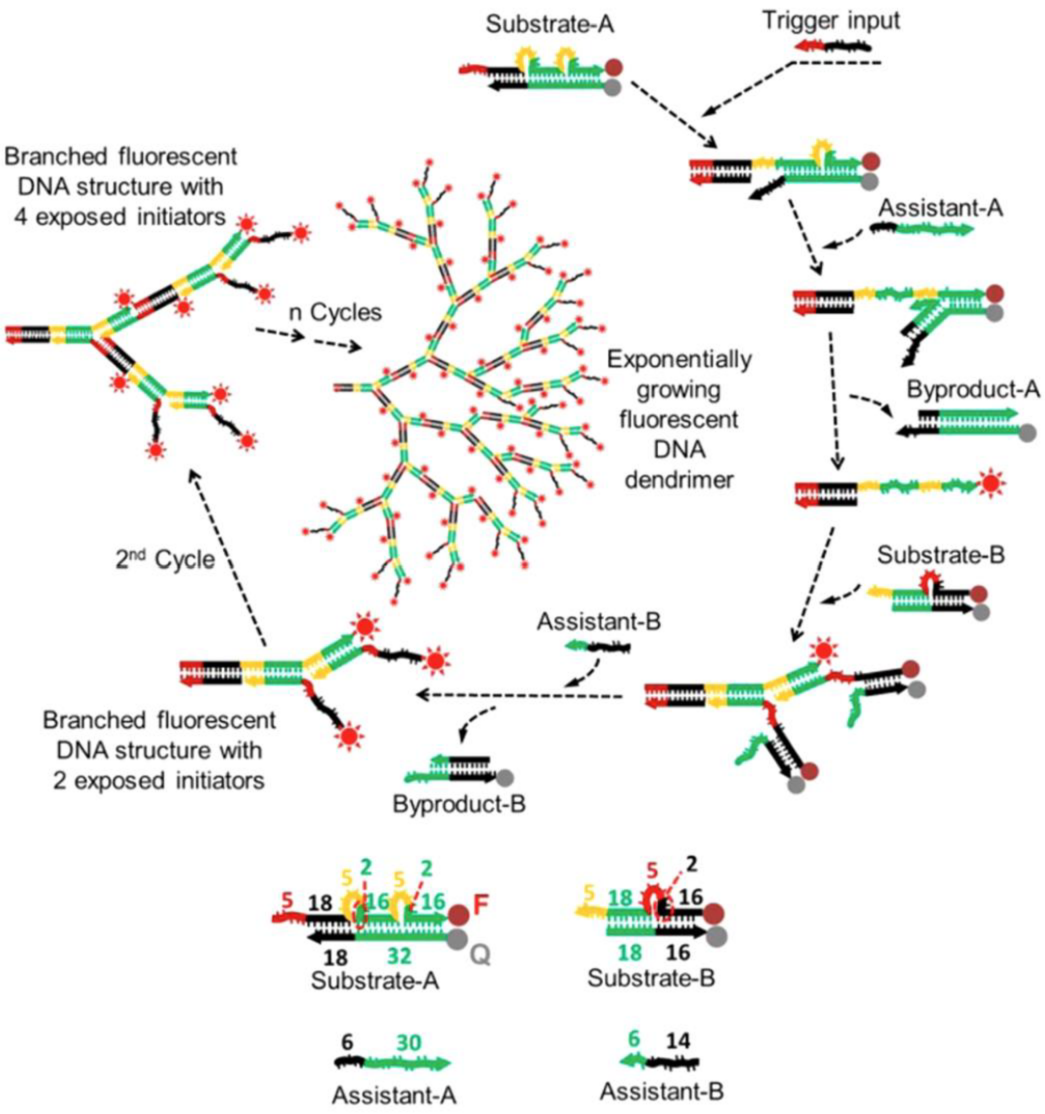


Figure 1.11 *Exponential growth of dendrimer in branched HCR.*

Instead of hairpins with single toehold loop, the substrate duplex A and B contains two and one toehold loops. Following initiation by a trigger, two assistant strands are required to further open A and B and ensure efficiency of the polymerization reaction. Reprinted with permission from ref 47. Copyright 2014 American Chemical Society.

1.4.4.2 DNA catalytic reactions using TMSD and TE

Turberfield et al.[40] developed the first non-covalent DNA catalytic reaction that involved both TMSD and TE (**Figure 1.12A)**. In this reaction, a loop complex **LS*** and a ssDNA **L***, which is complementary to **L** in **LS***, exist in a metastable state. Although **L*** binds to the loop region of **LS***, hybridization between **L** and **L*** is sterically hindered in the absence of **C** as **L*** has difficulties threading through the loop of **LS**. Therefore, the formation of the products, the duplex **LL*** and ssDNA **S***, is relatively slow. The catalytic DNA **C*** removes steric hinderance and expedites the reaction by opening the stem region of the complex through TMSD to liberate the loop region. The catalytic DNA is not consumed in this process as **C** is released through toehold exchange to facilitate further reaction. From a thermodynamic standpoint, the formation of a stable product from metastable or less stable reactant is energetically favorable. In this reaction, the additional base pair interaction in the loop region of **L** and **L*** also increases base-stacking, making the product **LL*** more stable, and the reaction is energetically favorable. Seelig et al.[30] modified the initial design to improve its performance, where a metastable complex **F** is designed to consist of two loop complexes, **LS*** and **L*S**, connected through a kissing interaction **(Figure 1.12B)**. The complex **F** remains stable for at least a week with little non-specific degradation into the products **LL*** and **SS***. The catalyst **C*** opens the loop complex **LS*** through TMSD and forms the complex **I1**. Further hybridization between **L** of the **LS*** loop and **L*** of **L*S** loop forms the Holiday junction that is **I2**, and releases **C***, which then catalyzes breakdown of another **F** complex. Through branch migration of four-way-junction **I2** eventually forms the products **LL*** and **SS***. Experiments showed the system achieved 40-fold turnover. The potential cause for this limited turnover is due to defective oligomers that traps the strands in the intermediate complexes. Conversely, other studies also point out defective oligos also can lead to non-specific background interactions that leads to background signals[49].

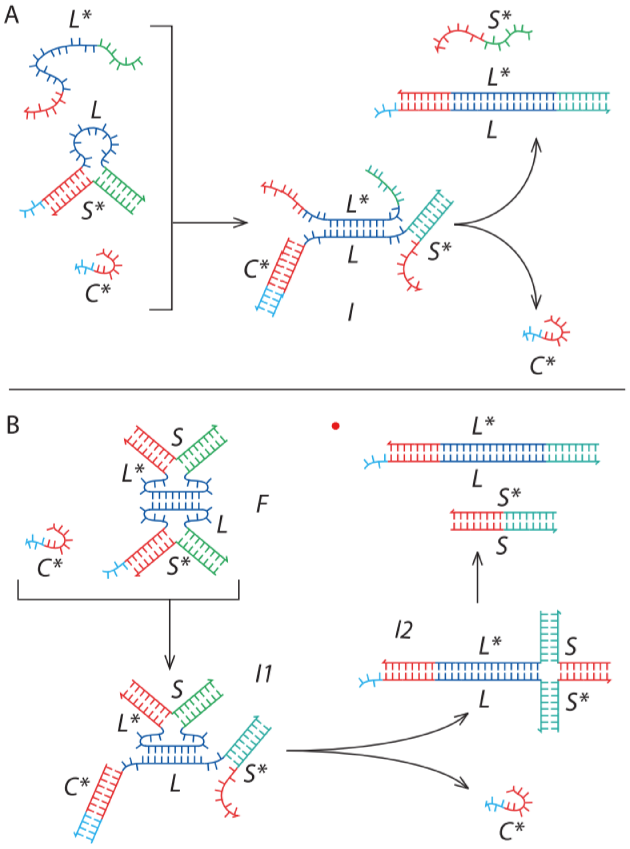


Figure 1.12 *Early non-covalent DNA catalytic reactions.*

(A) Scheme of Turberfield et al.[40] *Here, the reaction between the loop complex* ***LS**** *and single strands DNA* ***L**** *is expedited by the catalytic DNA* ***C****, *which enabled the formation of the intermediate* ***I*** *that quickly collapses into the products* ***S**** *and* ***LL****. *(B) Scheme of Seelig et al.*[30] *Here, the metastable complex* ***F*** *which consisted of the loop complexes* ***LS**** *and* ***L*S*** *connected through a kissing interaction is converted into the products* ***LL**** *and* ***SS**** *in the presence of the catalyst* ***C****. **C** facilities the formation of the intermediate* ***I1*** *and is released when* ***I2*** *is formed. The catalyst strand* ***C****enhances the rate of reaction by opening the stems of the loop complex* ***LS**** *through toehold mediate strand displacement****.*** *Reprinted with permission from ref 5. Copyright 2019 American Chemical Society.*

Another reaction that uses both TMSD and TE is entropy driven catalysis (EDC) developed by Zhang et al[50] (**Figure 1.13**). EDC is initiated by a TE reaction followed by a TMSD reaction and another TE reaction for regeneration of the catalytic strand (**Figure 1.13A).** The substrate complex **S** is formed by hybridization of **OB** with **SB** and **LB**. The catalyst **C** first react with the substrate **S** through a TE reaction, which exposes a new toehold $\overline{\mathbf{3}}$. Consequently, the domain **3** of the fuel DNA **F** reacts with the toehold $\overline{\mathbf{3}}$, initiating a TMSD and a TE reaction. Therefore, in addition to generation of a product **W**, the catalyst **C** is regenerated, and the output DNA **OB** is also displaced off. The **OB** then reacts with a reporter strand, generating fluorescence signals. The regenerated **C** then reacts with another substrate molecule, leading to cyclic catalytic reactions. Because there is no change in base pairs between the substrate molecule **S** and waster product **W**, there is little change in enthalpy for the overall reaction, where the substrate **S** reacts with the fuel **F**, generating the waste product **W** and two ssDNA strands **SB** and **OB**. However, because there are three products, while there are only two reactants, the entropy is increased for the overall reaction. Therefore, the reaction is called entropy-driven catalytic reaction. Simmel et al.[5] estimated the strength of this entropic driving force. The entropic driving force was found to be about 13.84 kcal/mol, which is approximately equivalent to 8 base pairs. The amplification efficiency of EDC is relatively low (~5 folds per hour)[33, 50, 51].

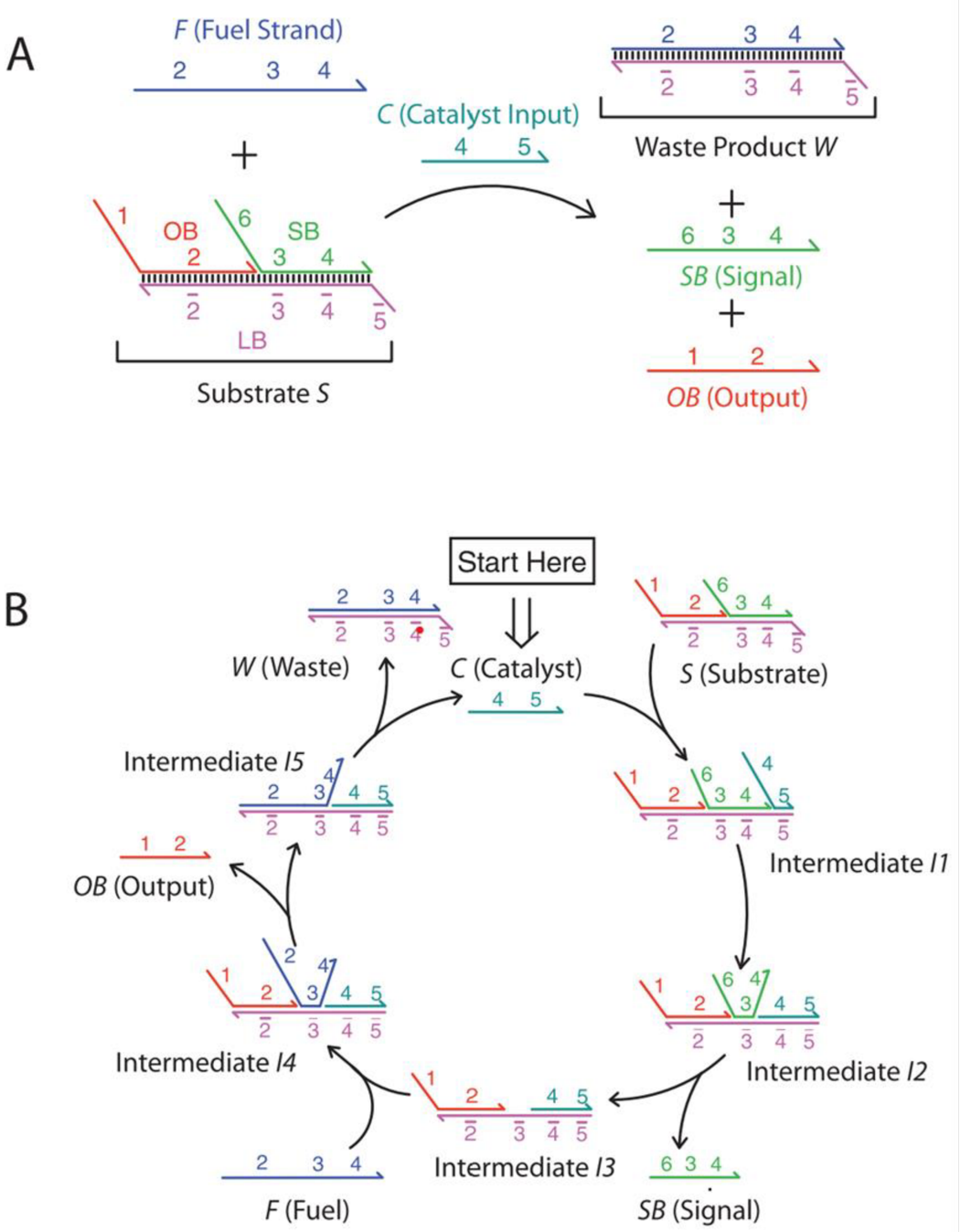


Figure 1.13 *An entropy-driven catalytic system for Non-covalent DNA amplification.*

The overall reaction is shown in (A). The two reactants, fuel strand ***F*** *and the substrate, generates the 3 products: waste product* ***W****, signal strand* ***SB****, and the output strand* ***OB****. This reaction is catalyzed by strand* ***C****. The catalytic process is depicted in (B), with the hypothesized intermediates displayed. The toeholds, domains* 3^{-} *and* 5^{-}*, are designed to be short with domain* 5^{-} *slightly more energetically stable compared to* 3^{-} *to ensure SB is easily displaced by the catalyst, similarly the catalyst is easily displaced by F. Reprinted with permission from ref 50. Copyright 2007 AAAS.*

Catalytic hairpin assembly[52, 53] (CHA) is another DNA catalytic reaction using TSMD and TE (**Figure 1.14**). The CHA consists of two hairpins **H1** and **H2**, where the loop of **H1** is fully or partially complementary to the stem of **H2** and vise versa. Although the reaction between **H1** and **H2** is energetically favorable, there are no complementary sequences that are free simultaneously. Therefore, the reaction between **H1** and **H2** is very slow. In the presence of the catalytic strand **C1**, it first reacts with one stem (containing domains **1**, **2**, and **3**) of **H1** through a TMSD, opening the hairpin **H1** to release the domain **3***, which is previously trapped in the stem region of **H1**. The domain **3*** then binds the toehold **3** of the hairpin, **H2**, initiating the reaction between **H1** and **H2**. Finally, **H2** can hybridize with **H1**, forming a duplex product. **H1** and **H2** does not completely hybridize with each other, and the displacement and regeneration of **C1** is achieve via toehold exchange. The **C1** then reacts with another **H1**, triggering cyclic reactions between **H1** and **H2**.

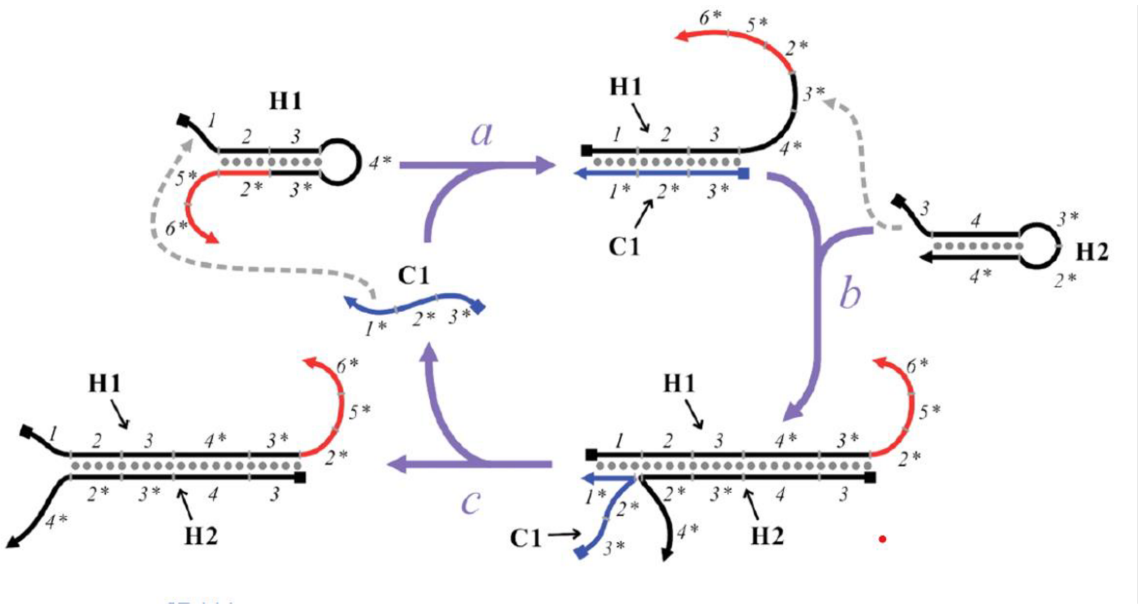


Figure 1.14 *Catalytic hairpin assembly (CHA) based on both TMSD and TE.*
*Strand **C1** catalyzes hybridization assembly of the substrate **H1** with the fuel **H2** and is regenerated at the end of each cycle. The exposed domain **2***, **5*** and **6*** on the product **H1·H2** duplexes can be used for fluorescence signal generation through TMSD reaction with DNA reporter. The CHA design can be used for sensor applications where the target analyte enters as **C1** respectively. Reprinted with permission from ref 53. Copyright 2011, Oxford University Press.*

1.4.4.3 DNA catalytic reaction based on TE.

Leveraging the reversible nature of TE for the regeneration of the catalytic strands, a cascade of two TE reactions are developed for cyclic amplification of DNA target[24] (**Figure1.15)**. The catalytic DNA, **C**, initiates the reaction by binding to the toehold domain **1** of the substrate **S** and initiate the TE reaction to displaces the output strand, **O**. The hybridization of **C** with the top strand of the **S** to form the intermediate, **I** with a new toehold, domain **4**. The second TE reaction regenerates **C** when the fuel, **F**, binds to the toehold on **I**, domain **4** to form a waste product and displaces **C**. The released **C** is free to react with another substrate to initiate another round of reaction to achieve cyclic amplification. Reaction between **O** and a TMSD reporter enable

fluorescence signal generation. Zhang et al.[24] used this type of reaction to establish the relationship between the length of toehold and the kinetics of both TMSD and TE.

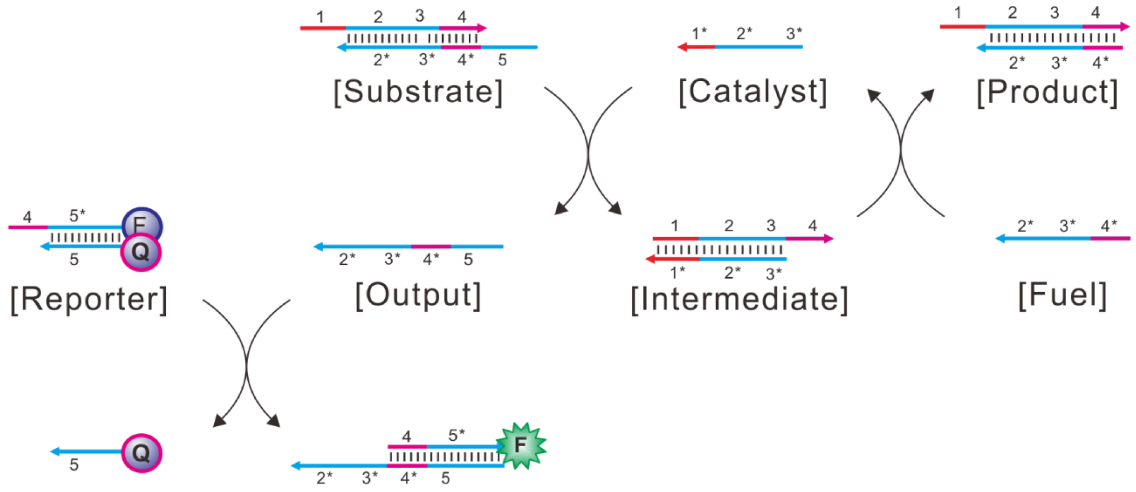


Figure 1.15 *Schematics of a non-covalent DNA catalytic reaction based on toehold exchange. The output of the toehold exchange reaction is consumed in a second toehold mediated strand displacement reaction. The output from the TE reaction enters the second reaction as a reactant. Through binding to the toehold domain 4 on the reporter, the output displaces the quencher to form a product with the fluorophore labelled strand. The fluorescence is then restored. The circle with Q represent the quencher while the circle with F at its center represent the quenched fluorophore.*

Knowing the importance of the length the forward and reverse toehold in TE with respect to k_f and k_r. In their work, Zhang and Winfree[24] set the length of the reverse toehold domain to be 5 nt long and investigated how the length of the forward toehold affects the overall reaction. They adjusted the length of the forward toehold to be anywhere from 2 nt to 9 nt. When the length of reverse toehold is shorter than the reverse toehold, such as when 1 = 2 nt, 3 nt, and 4 nt, or ΔG_f^o of 3:3* < ΔG_f^o of 1:1*. The ΔG_f^o of the substrate is lower than the ΔG_f^o of the intermediate, and the reaction between and **S** and **C** to form **I** and **O** is energetically unfavorable as well as having a slower kinetics compared to the reverse reaction. This first step reaction is found to be very slow

and is the rate limiting step of the overall reaction. The free energy change of this reaction would resemble **Figure, 1.16a**. When ΔG_f^o 3:3* =ΔG_f^o 1:1*, the free energy change of reaction is similar to **Figure 1.16b**. This reaction can be entirely driven by the concentration of Substrate, Fuel and Reporter, where small amount of the catalytic DNA is needed to initiate the reaction. While the equilibrium constant *K*, is defined as a ratio between k_f and k_r, the equilibrium concentration also depends on the initial concentration of the reactants and product. High concentration of reactant shifts the equilibrium position and force the formation of the products and vise versa. This design is also used for the development of see-saw gate and large-scale DNA circuits[54-56]. Lastly, when the length of forward toehold is longer than the reverse toehold, or ΔG_f^o 3:3* >ΔG_f^o 1:1*, the reaction is also slow. Although the first step reaction where the catalyst reacts with substrate to form the intermediate and output is fast in this situation, the second step reaction between the intermediate and fuel to form the waste and regenerate the catalyst is energetically unfavorable with slow reaction kinetics and is thus rate limiting. The free energy change of reaction would be similar to **Figure 1.16c**.

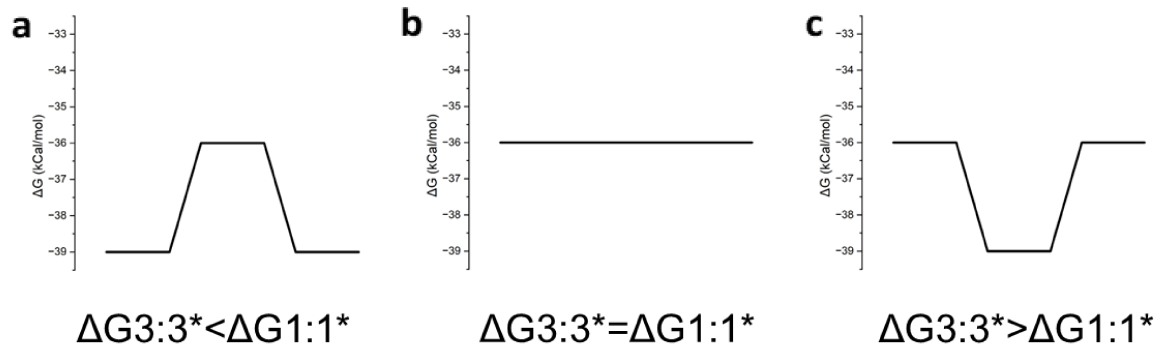


Figure 1.16 *Reaction energy diagram of the non-covalent DNA catalytic reaction with respect to stablity of the forward and reverse toeholds.*

a) When the reverse toehold is more stable compared to the forward toehold, the first partial reaction between the ***S*** *and* ***C*** *is not as energetically favorable as the reaction goes from a stable low energy substrate to an less stable higher energy intermediate. Therefore the overall reaction rate is limited by the the kinetic of the first partical reaction eventhough the second reaction is energtically. b) When the toeholds are equal in stablity, the reaction would qickly reach equliberium and remain constant. c) When the forward toehold is more stable compared to the reverse toehold, the second partical reaction becomes the rate limiting reaction. In all three scenarios the overall reaction free energy change* $\Delta \boldsymbol{G}_{OR}$ *is equal to 0 and the reaction is not energetically favorable. The values for* ΔG_f^o *are arrbituary values, and only meant to help visualize the free energy change of reaction.*

Besides controlling the length of the toehold, other method has been designed to improve the kinetic of the TE based reaction such as using DNA mismatches. DNA mismatches occur when the purine or pyrimidine base is not matched with its complementary base, such as when A is not matched with T or G is not matched with C or vise-versa. The mismatched bases can not form the same hydrogen bonds formed in complementary base pairs, and the structure of the DNA double helix is disrupted forming a bulge[57-59]. The two strands can become separated at the mismatch site and both the base pairing and base stacking interaction is disrupted. Thus, the stability of DNA hybridizations between the two strands is compromised and free energy of hybridization is increased[13, 57-59]. The impact of mismatch is highly dependent on the base that is mutated and its

complementary base, and the base that replaces the mutated base. For example, C is notoriously intolerant to mismatched bases on the complementary strand and can cause a big change in free-energy of hybridization while G is the most accommodating of mismatched bases and causes less change, and A and T fall somewhere in between. The presence of a mismatched base can be useful to speed up non-covalent DNA catalytic reaction as it can provide a hidden toehold or pit stop during branch migration to favor strand invasion while restrict the random walking of the reverse reaction.

While Broadwater and Kim[63] initially investigated the effect of DNA mismatches on TMSD, Haley and coworkers[59] performed a more systemic study of DNA mismatch on toehold exchange and non-covalent DNA catalytic reactions. Building on the design of Zhang and Winfree[24] (**Figure 1.15**), Haley et al. inserted C-A and T-T mismatches at the 12th and 11th position in the branch migration of substrate **BD**[58] (**Figure 1.17**). The mismatch is introduced by modification of the bases in **B**. However, it is important to note the DNA mismatches were inserted into the sequences both **A** and **B**, thus $\mathbf{A_{CA}}$ and $\mathbf{A_{TT}}$ is used to activate $\mathbf{B_{CA}D}$ and $\mathbf{B_{TT}D}$ respectively, which means the intermediate **AD** would also have the same mismatches as in **BD**. Therefore, the introduction of mismatches did not significantly improve the kinetics of the first step, where the catalytic strand **A** invade the substrate **BD** via the initial toehold (highlighted in red on **D**). Rather the insertion of mismatches improves the kinetics of the second reaction where the fuel strand **C** is used regenerate **A**. Haley et al. was able to show that the introduction of mismatch improved the over all reaction kinetics compared to a system without mismatch as well as when using an extended strand **C** with 2 extra base pairs at its 3' end. The uses of DNA mismatch had minimal increase in leakage unlike the design using the extended C strand. This demonstrates the efficacy of DNA mismatches as a way to design hidden toeholds that improved the efficiency of branch migrations.

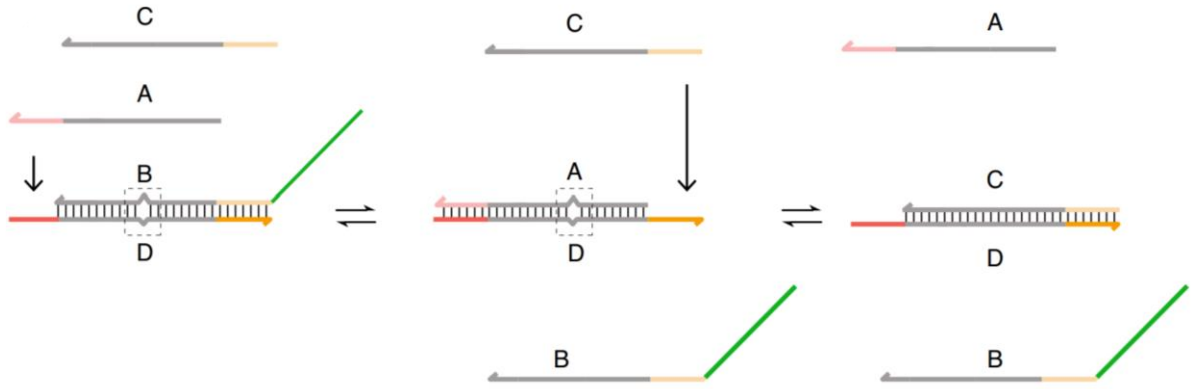


Figure 1.17 *Schematics of the non-covalent DNA catalytic reaction driven by mismatch elimination. The catalytic activator **A** reacts with the mismatched substrate **BD** to form the intermediate **AD** with the same mismatch. Recovery of **A** and elimination of the mismatch is achieved when the intermediate reacts with the fuel **C** to form the product **CD**. The recovered **A** can then initiated further round of rection to achieve cyclic amplification. Retention of the mismatch in product and intermediate means the kinetic of the first step reaction is unaffected, while the elimination of mismatch in the formation of the product from the intermediate improved the kinetic of the second reaction and in turn the overall reaction. Reused from ref. 59 under the CC BY license. Copyright © 2020, Natalie E. C. Haley et al.*

1.5 Objectives and rationales

Limitation of the current non-covalent DNA amplification reactions discussed in 1.4.4 include low amplification efficiency and complex design[33, 37, 48-50]. Specifically, the amplification efficiency of HCR and EDC is about 3-5 fold per hour, insufficient for detection of low levels of intracellular nucleic acid target such as miRNA[51, 52]. Although CHA has good amplification efficiency, the use of metastable hairpin structures as DNA substrates require thoughtful design of each functional domain, single-stranded overhang, hairpin stem and loop, to perform properly[41,48,59]. Additionally, chemical synthesis of these long hairpins, used in HCR and CHA, is

more prone to producing impurities which result in malformation of hairpins and then lead to leakage reactions, increasing background signals.

I aim to construct simple DNA catalytic reactions with improved amplification efficiency, building on the toehold exchange catalytic reaction reported by Zhang et al.[24] (**Figure 1.15**). The toehold exchange catalytic reaction is simple and only requires a substrate of dsDNA with a toehold overhang and a single-stranded fuel DNA. Destabilizing the substrate and raising its ΔG_f^o can make the forward reaction energetically favorable, allowing high signal amplification, while the kinetic of the non-specific leakage reaction is controlled by minimally changing the DNA base pairing for low background signal. The substrate can be destabilized by introducing DNA structures, mismatches, or chemically modified nucleotides into the output strand thus affecting its interaction with its complementary strand. Without destabilization, the free energy change of the overall reaction would be zero, and the reaction is completely governed by the principle of toehold exchange. However, depending on the length and stability of both the forward and reverse toeholds, the energy diagram of the reaction can show one of three different scenarios, as shown in **Figure 1.16**. The scenario 1(**Figure 1.16a**), where the forward toehold is less stable compared to the reverse toehold and first partial reaction is non-spontaneous and rate limiting is considered. Modifications to the output strand serve as several purposes. First, it would raise the free energy of hybridization of the substrate (ΔG_{fs}^o) to make the overall reaction energetically favorable (**Figure 1.18**). Secondly, raising the ΔG_{fs}^o could help the make the first step reaction or the formation of intermediate more energetically favorable. Finally, the change in ΔG_{fs}^o and modification of the substrate would impact the kinetics of either the branch migration or the dissociation of the reverse toehold following the branch migration process, and thus preventing the reverse reaction. Therefore, the modification of the output strand would help make the non-

covalent DNA catalytic reaction much more energetically favorable and improve the amplification efficiency.

Herein, I introduce DNA nanostructure, DNA mismatches, and chemically modified nucleotides into the output strand and substrate to modulate ΔG^o_{fs}. In chapter 2, I use the three-way-junction (TWJ) to create tension at the junction and destabilize the base-stacking interaction,[13, 61, 62] which causes large changes in ΔG^o_{fs}, making the reactions more thermodynamically favorable.

In chapter 3, I introduce DNA mismatches into the output strand to destabilize the substrate. There are two advantages of using mismatch compared to using TWJ: 1) the structure of substrate is simpler since the substrate of the mismatch-aided catalytic reaction is a simple double-stranded duplex containing a designed mismatch, whereas the substrate of TWJ-mediated catalytic reaction needs three strands, or the output strand needs to contain a stem-loop; 2) the use of mismatch allows me to readily modulate the stability of the substrate.

Finally, in chapter 4, I utilize modified nucleotides to destabilize the substrate. The three main constituents of nucleic acids: the base, pentose sugar, and the phosphate linkages can all be modified[63,64]. Modified nucleotides are introduced into the oligonucleotide during in vitro synthesis. The modified nucleotides can have different chemical and physical properties that either stabilize or destabilize the hybridization of the DNA double helix[63, 64]. These modified nucleotide does not disrupt the secondary structure of DNA in the same way as the DNA mismatch, and this can lead to lower leakage. Although Young & Sczepanski incorporated PNA for the recognition of L-formed DNA[65], Olson et al.[66] explored the effects of LNA in strand displacement, and Wang and coworkers leveraged abasic site for intracellular activation of non-covalent DNA catalytic reaction and detect of endonucleases activity[67, 68], there is a need to investigate how modified

nucleotides destabilize the DNA duplex and improve the thermodynamics and kinetics of non-covalent DNA catalytic reaction.

The development of DNA catalytic reactions through the site specific, strategic incorporation of motifs such as TWJ, mismatches, and modified nucleotide has important potentials for intracellular detection, point-of-care testing, targeted drug delivery, and dynamic molecular circuits. For example, these techniques could be applied to studying the dynamic functions and profiles of nucleic acids, such as miRNA in living cells. The DNA catalytic reaction can operate isothermally at a range of temperatures, including room temperature and 37°C, without the need for any enzymes, which is ideal for intracellular application. In addition to amplified detection of intracellular miRNA, the DNA catalytic reactions can be utilized to achieve targeted drug delivery and improve selective drug targeting and reduce side effects. Further amplification could be achieved through the construction of exponential amplification, where the released output strand can also serve as the catalyst to react with substrates and fuel to initiate more catalytic reactions.

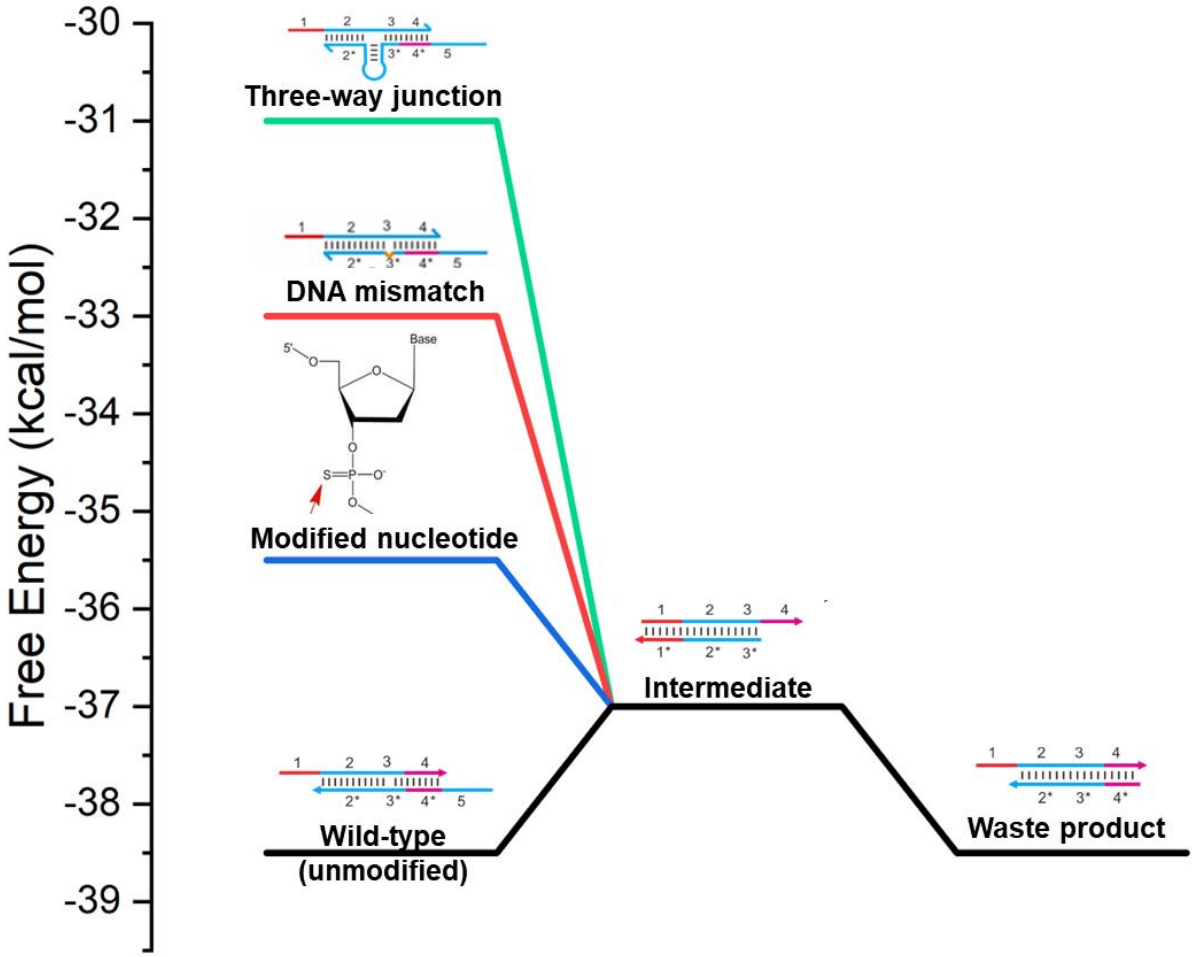


Figure 1.18 *Expected change in reaction thermodynamic and kinetic changes in response to introduction of modification output strand.*

Modification of the output strand would affect its binding with the complementary strand, and destablize the substrate, thus raising the ΔG^{o}_{fs}*. As a result, the free energy changes of both the first elementary step and the overall reaction would be negative and the reactions are spontaneous. When the modification is introduced into the output strand, the interaction between the output and the complementary strand is destablized but the interaction between the catalytic strand and and the complemenetary strand is unchanged. Therefore, the branch migration in the forward direction is facilitated and the reverse reaction is negligible in comparison.*

1.6 References

1. Watson, J. D., & Crick, F. H. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature. 171(4356), 737.
2. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Molecular Biology of the Cell (4th ed.). Garland Science.
3. Franklin, R. E., & Gosling, R. G. (1953). Molecular Configuration in Sodium Thymonucleate. Nature. 171(4356), 740.
4. Crick, F. H., Barnett, L., Brenner, S., & Watts-Tobin, R. J. (1961). General Nature of the Genetic Code for Proteins. Nature. 192(4809), 1227.
5. Simmel, F. C., Yurke, B., & Singh, H. R. (2019). Principles and Applications of Nucleic Acid Strand Displacement Reactions. Chem Rev. 119(10), 6326.
6. Seeman, N. C., & Sleiman, H. F. (2017). DNA nanotechnology. Nat Rev Mater. 3(1). 17068.
7. Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., & Hofreiter, M. (2004). Genetic analyses from ancient DNA. Annu Rev Genet. 38, 645.
8. Keller, A., & Linko, V. (2020). Challenges and Perspectives of DNA Nanostructures in Biomedicine. Angew Chem Int Ed. 59(37), 15818.
9. Arora, A. A., & de Silva, C. (2018). Beyond the smiley face: applications of structural DNA nanotechnology. Nano Rev Exp. 9, 1430976.
10. Ke, Y., Castro, C., & Choi, J. H. (2018). Structural DNA nanotechnology: artificial nanostructures for biomedical research. Annu Rev Biomed Eng. 20, 375.
11. Beaucage, S. L., & Caruthers, M. H. (1981). Solid-phase synthesis of oligodeoxyribonucleotides by the phosphoramidite method. Tetrahedron Lett. 22(20), 1855.
12. Caruthers, M. H. (1985). Chemical synthesis of DNA and RNA: An overview. Bioorg Chem. 13(3), 217.
13. SantaLucia Jr., J., Hicks, D. (2004). The thermodynamics of DNA structural motifs. Annu Rev Biophys. 33, 415.
14. SantaLucia Jr., J. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci U S A. 95, 1460.

15. Dirks, R. M., Bois, J. S., Schaeffer, J. M., Winfree, E., & Pierce, N. A. (2007). Thermodynamic analysis of interacting nucleic acid strands1. SIAM Review. 49(1), 65.
16. Zhang, J. X., Fang, J. Z., Duan, W., Wu, L. R., Zhang, A. W., Dalchau, N., Yordanov, B., Petersen, R., Phillips, A., Zhang, D. Y. (2017). Predicting DNA hybridization kinetics from sequence. Nat Chem. 10, 91.
17. Wetmur, J. G., & Davidson, N. (1968). Kinetics of renaturation of DNA. J Mol Biol. 31, 349.
18. Yurke, B. & Mills, A. P., (2003). Using DNA to power nanostructures. Genet Program Evol M. 4, 111
19. Morrison, L. E., & Stols, L. M. (1993). Sensitive fluorescence-based thermodynamic and kinetic measurements of DNA hybridization in solution. Biochem. 32(12), 3095.
20. Reynaldo, L. P., Vologodskii, A. V., Neri, B. P., Lyamichev, V. I. (2000). The kinetics of oligonucleotide replacements. J Mol Biol. 297(2), 511.
21. Zhang, J. X., Fang, J. Z., Duan, W., Wu, L. R., Zhang, A. W., Dalchau, N., Yordanov, B., Petersen, R., Phillips, A., & Zhang, D. Y. (2017). Predicting DNA hybridization kinetics from sequence. Nature. 6(11), 2877.
22. Zhang, D. Y. (2010). Dynamic DNA Strand Displacement Circuits. California Institute of Technology. Pasadena, California
23. Green, C., & Tibbetts, C. (1981). Reassociation rate limited displacement of DNA strands by branch migration. Nucleic Acids Res. 9(8), 1905.
24. Zhang, D. Y., Winfree, E. (2009). Control of DNA strand displacement kinetics using toehold exchange. J Am Chem Soc. 131, 17303.
25. Srinivas, N., Ouldridge, T. E., Sulc, P., Schaeffer, J. M., Yurke, B., Louis, A. A., Doye, J. P. K., & Winfree, E. (2013). On the biophysics and kinetics of toehold-mediated DNA strand displacement. Nucleic Acids Res. 41(22), 10641.
26. Ouldridge, T. E., Louis, A. A., Doye, J. P. K. (2010). DNA nanotweezers studied with a coarse-grained model of DNA. Phys Rev Lett. 104, 178101.
27. Ouldridge, T. E., Louis, A. A., Doye, J. P. K. (2011). Structural, mechanical and thermodynamic properties of a coarse-grained DNA model. J Chem Phys. 134, No. 085101.
28. Ouldridge, T. E. Coarse-grained modeling of DNA and DNA nanotechnology. Ph.D. Thesis, University of Oxford, 2011

29. Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., Pierce, N. A. (2011). NUPACK: analysis and design of nucleic acid systems. J Comput Chem. 32, 170.
30. Seelig, G., Yurke, B., & Winfree, E. (2006). Catalyzed relaxation of a metastable DNA fuel. J Am Chem Soc. 128(30), 12211.
31. Kotani, S., & Hughes, W. L. (2017). Multi-arm junctions for dynamic DNA nanotechnology. J Am Chem Soc. 139, 6363.
32. Panyutin, I. G.; Hsieh, P. (1994). The kinetics of spontaneous DNA branch migration. Proc Natl Acad Sci U S A. 91, 2021.
33. Peng, H., Newbigging, A. M., Wang, Z., Tao, J., Deng, W., Le, X. C., & Zhang, H. (2018). DNAzyme-Mediated Assays for Amplified Detection of Nucleic Acids and Proteins. Anal Chem. 90(1), 190.
34. Yurke, B., Turberfield, A. J., Mills Jr, A. P., Simmel, F. C., & Neumann, J. L. (2000). A DNA-fuelled molecular machine made of DNA. Nature. 406(6796), 605.
35. Yin, P., Yan, H., Daniell, X. G., Turberfield, A. J., & Reif, J. H. (2004). A unidirectional DNA walker that moves autonomously along a track. Angew Chem Int Ed. 43(30), 4906.
36. Shin, J.S., & Pierce, N. A. (2004). A synthetic DNA walker for molecular transport. J Am Chem Soc. 126(35), 10834.
37. Zhang, D. Y., & Seelig, G. (2011). Dynamic DNA nanotechnology using strand-displacement reactions. Nat Chem. 3(2), 103.
38. George, A. K., & Singh, H. (2016). Enzyme-free scalable DNA digital design techniques: a review. IEEE Trans Nanobioscience. 15(8), 928.
39. Seelig, G., Soloveichik, D., Zhang, D. Y., & Winfree, E. (2006). Enzyme-Free Nucleic Acid Logic Circuits. Science. 314(5805), 1585.
40. Turberfield, A. J., Mitchell, J. C., Yurke, B., Mills, A. P., Blakey, M. I., & Simmel, F. C. (2003). DNA fuel for free-running nanomachines. Phys Review Lett. 90(11), 118102.
41. Green, S. J., Lubrich, D., & Turberfield, A. J. (2006). DNA hairpins: fuel for autonomous DNA devices. Biophys J. 91(8), 2966.
42. Dirks, R. M., Pierce, N. A. (2004). Triggered amplification by hybridization chain reaction. Proc Natl Acad Sci USA. 101, 15275.

43. Ang, Y. S., & Yung, L. Y. L. (2016). Rational design of hybridization chain reaction monomers for robust signal amplification. Chem Commun. 52(22), 4219.
44. Venkataraman, S., Dirks, R. M., Rothemund, P. W. K., Winfree, E., & Pierce, N. A. (2007). An autonomous polymerization motor powered by DNA hybridization. Nat Nanotechnol. 2(8), 490.
45. Cangialosi, A., Yoon, C., Liu, J., Huang, Q., Guo, J., Nguyen, T. D., Gracias, D. H., & Schulman, R. (2017). DNA sequence-directed shape change of photopatterned hydrogels via high-degree swelling. Science. 357, 1126.
46. Fern, J., & Schulman, R. (2018). Modular DNA strand-displacement controllers for directing material expansion. Nat Commun. 9, 3766.
47. Xuan, F., & Hsing, I. M. (2014). Triggering hairpin-free chain-branching growth of fluorescent DNA dendrimers for nonlinear hybridization chain reaction. J Am Chem Soc. 136(27), 9810.
48. Bi, S., Chen, M., Jia, X., Dong, Y., & Wang, Z. (2015). Hyperbranched hybridization chain reaction for triggered signal amplification and concatenated logic circuits. Angew Chem Int Ed. 127, 8262.
49. Chen, X., Briggs, N., McLain, J., & Ellington, A. D. (2013). Stacking Nonenzymatic Circuits for High Signal Gain. . Proc Natl Acad Sci USA, 110, 5386.
50. Zhang, D. Y., Turberfield, A. J., Yurke, B., Winfree, E. (2007). Engineering entropy-driven reactions and networks catalyzed by DNA. Science. 318, 1121.
51. Jung, C., & Ellington, A. D. (2014). Diagnostic Applications of Nucleic Acid Circuits. Acc Chem Res. 47(6), 1825.
52. Yin, P., Choi, H. M. T., Calvert, C. R., Pierce, N. A. (2008). Programming biomolecular self-assembly pathways. Nature. 451, 318.
53. Li, B., Ellington, A. D., & Chen, X. (2011). Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. Nucleic Acids Res. 39(15), e110
54. Qian, L., & Winfree, E. (2011). Scaling up digital circuit computation with DNA strand displacement cascades. Science. 332(6034), 1196.
55. Qian, L., & Winfree, E. (2011). A simple DNA gate motif for synthesizing large-scale circuits. J R Soc Interface. 8, 1281.

56. Qian, L., Winfree, E., & Bruck, J. (2011). Neural network computation with DNA strand displacement cascades. Nature. 475, 368.
57. Rossetti, G., Dans, P. D., Gomez-Pinto, I., Ivani, I., Gonzalez, C., & Orozco, M. (2015). The structural impact of DNA mismatches. Nucleic Acids Res. 43(8), 4309.
58. Broadwater, D. W., & Kim, H. D. (2012). The effect of basepair mismatch on DNA strand displacement. Biophys J. 102(12), 2915.
59. Haley, N. E. C., Ouldridge, T. E., Mullor Ruiz, I., Geraldini, A., Louis, A. A., Bath, J., & Turberfield, A. J. (2020). Design of hidden thermodynamic driving for non-equilibrium systems via mismatch elimination during DNA strand displacement. Nat Commun. 11(1), 2461.
60. Jiang, Y. S., Li, B., Milligan, J. N., Bhadra, S., Ellington, A. D. (2013). Real-time detection of isothermal amplification reactions with thermostable catalytic hairpin assembly. J Am Chem Soc. 135, 7430.
61. Ladbury, J. E., Sturtevant, J. M., & Leontis, N. B. (1994). The Thermodynamics of Formation of a Three-Strand, DNA Three-Way Junction Complex. Biochem. 33(22), 6828.
62. Carr, C. E., & Marky, L. A. (2017). Investigation of the Melting Behavior of DNA Three-Way Junctions in the Closed and Open States. Biophys J. 113(3), 529.
63. Duffy, K., Arangundy-Franklin, S., & Holliger, P. (2020). Modified nucleic acids: replication, evolution, and next-generation therapeutics. BMC Biology. 18, 1122.
64. Jafari, D., Shajari, S., & Samadikuchaksaraei, A. (2020). Modified nucleic acids: Expanding the capabilities of functional biomaterials. MDPI Molecules. 25(20), 46594.
65. Young, B. E., & Sczepanski, J. T. (2019). Heterochiral DNA strand-displacement based on chimeric D/L-oligonucleotides. ACS Synth Biol. 8(12), 2756.
66. Olson, X., Kotani, S., Yurke, B., Graugnard, E., & Hughes, W. L. (2017). Kinetics of DNA strand displacement systems with locked nucleic acids. J Phys Chem B. 121(11), 2594.
67. Chen, Y., Gong, X., Gao, Y., Shang, Y., Shang, J., Yu, S., Li, R., He, S., Liu, X., & Wang, F. (2021). Bioorthogonal regulation of DNA circuits for smart intracellular microRNA imaging. Chem Sci. 12(47), 15710.
68. Shang, Y., Chen, Y., Wang, Q., He, Y., He, S., Yu, S., Liu, X., & Wang, F. (2022). Construction of an endogenously activated catalytic DNA circuit for highly robust in vivo microRNA imaging. Nano Today. 45, 101553.

2. Chapter 2: Three-way-junction (TWJ)-mediated DNA catalytic reaction

2.1 Introduction

Non-covalent DNA catalytic reactions represent a class of molecular amplification techniques in the field of dynamic DNA nanotechnology[1-3]. These reactions involve the cyclic reaction of the catalyst DNA with DNA substrates through toehold-mediated strand displacement reactions, producing multiple copies of the DNA products from a single copy of the catalyst DNA. DNA substrates are intentionally designed to possess kinetically trapped structures, generating a high activation energy[1-3]. Consequently, the uncatalyzed reaction rate is slow, and the DNA substrates remain stable in the absence of the catalyst. In the presence of the catalyst, it reacts with the DNA substrate, disrupting the stable secondary structure of the DNA substrate and lowering the activation energy. Therefore, the reaction kinetics is improved. Similar to a catalyst in a chemical reaction, the DNA catalyst actively participates in the reaction but remains unconsumed, allowing the generation of multiple product molecules from a single DNA catalyst strand.

Since the advent of the first DNA catalytic reaction[3-5], three major types of DNA catalytic reactions have been developed and widely used for detection of nucleic acids and other molecules. These include hybridization chain reaction (HCR)[6-8], entropy-driven catalytic reaction (EDC)[9, 10], and catalytic hairpin assembly (CHA)[7, 11]. However, these catalytic reactions have limitations that impact their applications to biosensing. For instance, the amplification efficiency is typically about 2-5 fold per hour for HCR and EDC. HCR and CHA involves the substrates in hairpin structures, which are difficult to design and synthesize. Purification, such as by polyacrylamide gel electrophoresis (PAGE), is often needed to prepare high-quality substrates[1-3].

To overcome these limitations, the objective of this chapter was to develop a simple and efficient DNA catalytic reaction. The toehold exchange catalytic reaction, which is simple and only requires a substrate formed by hybridization of two ssDNA strands and a single-stranded fuel[10], was used to develop the catalytic reaction. The hypothesis was that the kinetics of the catalytic reaction could be enhanced by increasing the free energy of the substrate. To achieve this, the substrate was designed to form a three-way-junction (TWJ)[12-14] and its performance in improving the amplification efficiency of the catalytic reaction while maintaining low background due to leakage reactions was studied.

2.2 Experimental section

2.2.1 DNA oligonucleotides

DNA oligonucleotides (oligos) were synthesized, purified, and labeled by Integrated DNA Technologies (IDT, Coralville, IA). The sequence and labelling of the used DNA oligos are listed in Tables 2.1.

2.2.2 Substrate preparation

Substrates were prepared by simple annealing without the need for purification. For the preparation of the 3S-TWJ substrate, **D** was mixed with **O-1** and **O-2** at a molar ratio of 1:1.2:1.2 in 10 mM TE buffer (pH 8.0) containing 14 mM $MgCl_2$ and 0.05% w/v Tween 20, which is a surfactant reducing the adsorption of DNA, with a final concentration of **D** at 5 μM. For the preparation of the 2S-TWJ substrate, **S-O** was mixed with **D-1** at a molar ratio of 1:1.2 in 10 mM TE buffer (pH 8.0) containing 14 mM $MgCl_2$ and 0.05% w/v Tween 20, with a final concentration

of **S-O** at 5 μM. The mixtures were loaded into a thermocycler for annealing using a program of 80°C for 5 minutes, cooling down to 20°C at a rate of 0.1°C every 5 seconds. The substrates were stored at 4°C prior to use.

2.2.3 Free energy of formation estimation

The free energy of formation of the substrates, intermediates, and product DNA was estimated through NUPACK3[15-16]. The concentrations of Na^+ and Mg^{2+} were set at 50 mM and 20 mM, respectively. The free energy of the 3S-TWJ substrate was estimated through the equation: the free energy of the 3S-TWJ substrate = the binding energy of the TWJ substrate – the binding energy of **O-1** and **O-2**. Similarly, the free energy of the 2S-TWJ substrate was estimated through the equation: the free energy of the 2S-TWJ substrate = the binding energy of the TWJ substrate – the binding energy of the stem in **S-O**. The free energy of all single-stranded strands is considered zero because there are not stable secondary structures in those strands.

2.2.4 Polyacrylamide gel electrophoresis analysis

Polyacrylamide gel electrophoresis (PAGE) was performed using a 10% w/v PAGE gel. A 20 μL aliquot of the reaction solution was loaded into each well. The gel was run at 80 V/cm until the dye band reached approximately 3/4 of the gel length. The gel was then stained with SYBR Gold and imaged using the ImageQuant LAS 400 (GE Healthcare, USA). The intensities of the bands were measured using ImageJ[17].

2.2.5 Determination of rate constants

Rate constants k_1 and k_2 were determined through the following equations using elementary step 1 as an example.

$$\text{Rate} = -\frac{d[S]}{dt} = -\frac{d[C]}{dt} = \frac{d[I3]}{dt} = \frac{d[O]}{dt} = k_1[\text{S}][\text{C}] \qquad (1)$$

Where [S] is molar concentration of the substrate, [C] is molar concentration of the catalyst, and k_1 is the rate constant.

$$[S] = ([S]_0 - x); [C] = ([C]_0 - x). \qquad (2)$$

Where $[\text{S}]_0$ represents the initial concentration of the substrate, $[\text{C}]_0$ represents the initial concentration of the catalyst, and x represents amount of the substrate or catalyst that has been reacted. After substituting equation 2 into equation 1, the following equations are obtained:

$$\frac{d([S]_0 - x)}{dt} = -k_1([S]_0 - x)([C]_0 - x) \qquad (3)$$

$$-\frac{dx}{dt} = -k_1([S]_0 - x)([C]_0 - x) \qquad (4)$$

Solving of equation 4 was performed using MATLAB codes created by Zhang and Winfree[10] with minor modifications.

For accurate determination of the reaction process, 1 nM substrate and 2 nM catalyst were used in a 100µL reaction mixture. The reaction buffer consisted of 10 mM Tris-EDTA (TE) buffer (pH 8.0), 20 mM $MgCl_2$, and 0.05% w/v Tween 20. These relatively low reactant concentrations allowed the reaction to occur at a slow rate while generating a sufficient fluorescence signal for detection. To minimize the reverse reaction, a DNA translator was added to react with **O-2** of the output DNA, sequestering the output DNA. The reaction progress was monitored in real time using a fluorescence plate reader, with fluorescence detected every 3 minutes for one hour at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

Because the rate of the uncatalyzed reaction is slow, 100 nM substrate and 100 nM fuel were used in a 100μL-reaction mixture for determining k_0. Fluorescence was detected every 5 minutes for two hours. The reaction curves were then used to calculate 1/[S].

$$\frac{1}{[S]} = \frac{1}{[S]_0} + k_0 t \qquad (5)$$

Using equation 5, k_0 can be obtained by plotting 1/[S] against t.

2.2.6 Performing the TWJ-mediated DNA catalytic reaction

Unless stated otherwise, the TWJ-mediated DNA catalytic reaction was conducted using 100μL of reaction solutions containing 25 nM substrate, 100 nM fuel, and the desired concentration of catalyst in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. The reaction progress was monitored in real-time using a fluorescence plate reader, with fluorescence detected every 3 minutes for one hour at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

2.2.7 Performing the two layer TWJ-mediated DNA catalytic reaction

Unless stated otherwise, the two-layer TWJ-mediated DNA catalytic reaction was conducted using 100μL of reaction solutions containing 25 nM layer1 substrate, 100 nM layer1 fuel, 200 nM translator, 20 nM layer 2 substrate, 200 nM layer 2 fuel, and the desired concentration of catalyst in 10 mM Tris-EDTA buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween. The reaction progress was monitored in real-time using a fluorescence plate reader, with fluorescence detected every 3 minutes for three hours at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

2.3 Results and discussion

2.3.1 Working principles of TWJ mediated DNA catalytic reaction

I chose to utilize the toehold exchange reaction to develop the TWJ-mediated catalytic reaction because it is simple and only requires a substrate of dsDNA with a toehold overhang and a single-stranded fuel DNA. As shown in **Figure 2.1 A** and **Figure 2.2**, the substrate is formed by the hybridization of **D** with **O-1'**. To simplify the explanation, I mapped each DNA strand into different functional domains labeled by numbers, where domain **n** is complementary to domain **n***. The arrow denotes the 3'-end of a DNA strand. Because the fuel DNA has the same sequence as **O-1'**, the reaction between the substrate and the fuel DNA generates two products identical to the substrate and fuel DNA, respectively. Therefore, the free energy change ΔG^0 of the reaction, ΔG^o_{OR}, is zero, which is not favorable for the catalytic reaction. Consequently, the amplification efficiency of this catalytic reaction is low, usually 2-5 fold per hour.

I hypothesize that the kinetics of the catalytic reaction can be enhanced by making the overall reaction energetically favorable. Since the free energy of single-stranded fuel DNA and output DNA is considered zero, the appropriate approach is to make the substrate less stable than the product DNA, thereby making the free energy of the overall reaction less than zero. To achieve this, I introduced a junction into the substrate to form a TWJ for two main reasons. First, the presence of the junction disrupts the normal base-stacking of dsDNA at the junction site, destabilizing the DNA hybridization of **D** and output DNA **O** (**Figure 2.1 B**, **2.2)**. Second, the introduction of the junction does not affect the base-pairing between the **D** and output DNA. Consequently, no nucleotides in the output DNA are unpaired, reducing the background due to nonspecific reactions between the substrate and fuel DNA.

Figure 2.1 B and **Figure 2.2** shows the design of the TWJ-mediated DNA catalytic reaction. The substate (**S**) has a TWJ formed by hybridization of **D** and the output DNA that is further formed by hybridization between **O-1** and **O-2**. To exclude effect of secondary structures on the kinetics of the reaction, I designed the **D** not to contain any stable secondary structures. In addition to the TWJ, the substate also contains a toehold (domain **1**). For real-time detection of the catalytic reaction, the 3'-end of **D** is labeled with a fluorophore and the 5'-end of **O-2** with a quencher. The formation of the TWJ places the fluorophore in close proximity to the quencher. The fuel DNA (**F**) is a single-stranded strand sharing the same domains **2***, **3***, and **4*** same with the output DNA. In the absence of the catalyst DNA (**C**), the reaction rate of the fuel DNA with the substrate is very slow because no nucleotides within domains **2***, **3***, and **4*** are exposed to serve as the toehold for increasing the reaction kinetics. This uncatalyzed reaction, also referred to as leakage, generates background signals.

In the presence of **C**, the catalyst can catalyze the reaction between the substrate and fuel strand through a multistep reaction involving two toehold exchange reactions (**Figure 2.1 C**). Specifically, the domain **1*** of **C** first hybridizes to the toehold **1** of the substrate, forming the first intermediate ***I1***. As a consequence, the first toehold exchange reaction is initiated, and through branch migration forms the second intermediate ***I2***. The toehold exchange reaction is completed when the output DNA dissociates from **D**, forming the third intermediate ***I3***. The formation of ***I3*** exposes the previously-blocked toehold domain **4**. Subsequently, **4*** of the fuel DNA binds with domain **4** of ***I3***, initiating the second toehold exchange reaction and generating the fourth intermediate ***I4***. The branch migration then forms the fifth intermediate ***I5*** from ***I4***. The **C** finally dissociates from **D**, forming the product DNA. Consequently, **C** is recycled and reacts with another substrate molecule. The process repeats, leading to linear signal amplification of **C**.

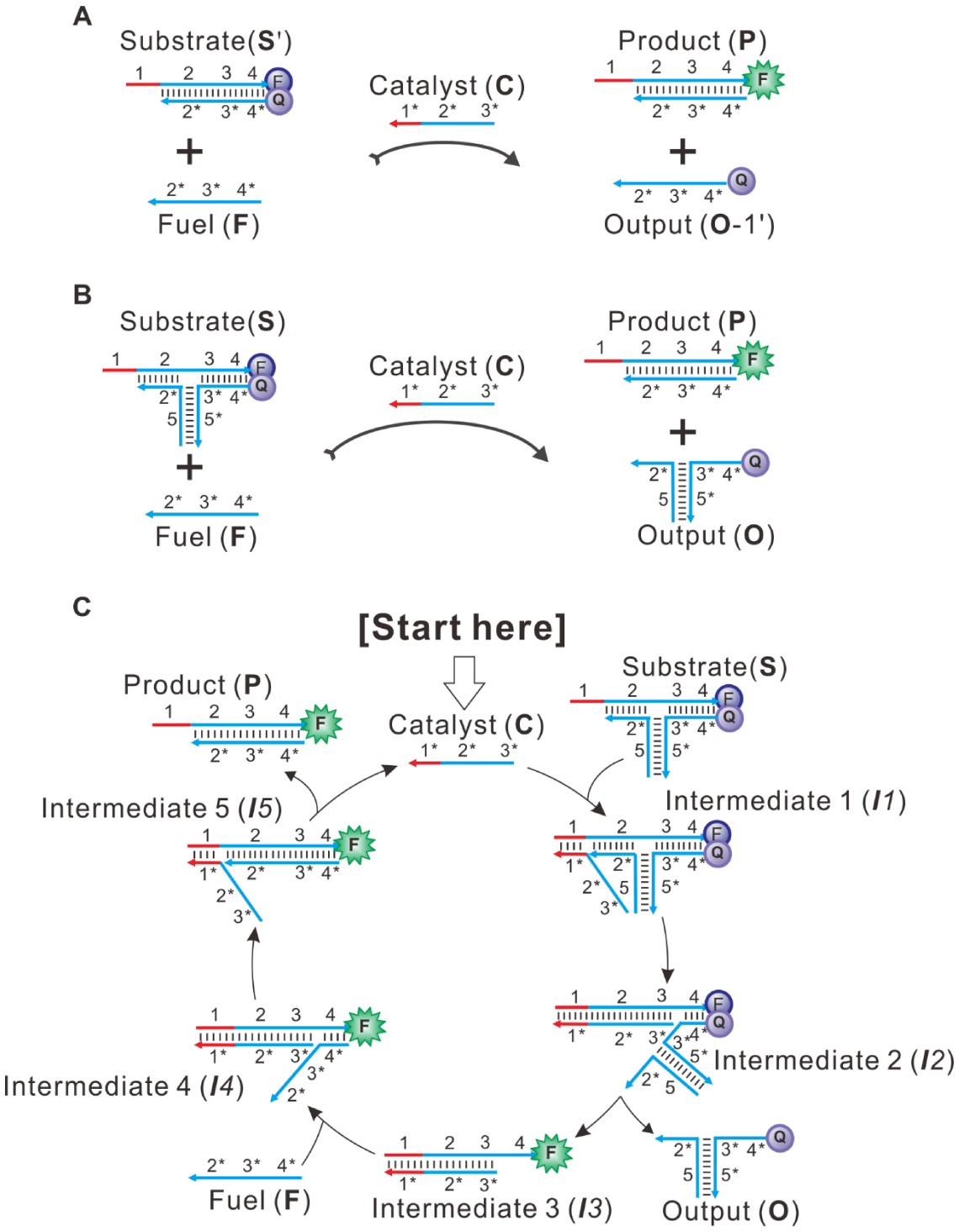


***Figure 2.1**: The three way junction-mediated catalytic reaction.*

*(**A**) Toehold exchange DNA catalytic reaction without TWJ. (**B**) TWJ-mediated DNA catalytic reaction. (**C**) The proposed reaction pathway of the TWJ-mediated DNA catalytic reaction.*

Because intermediates ***I1, I2, I4, I5*** are not stable, I describe this catalytic reaction using a simplified model. As shown by the following equations, two toehold exchange reactions are used as the elementary steps:

$$\text{Overall Reaction (OR): } S + F \underset{k_{-1}k_{-2}}{\overset{k_1k_2}{\rightleftarrows}} P + O$$

$$\text{Elementary reaction 1(es1): } S + C \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} I3 + O$$

$$\text{Elementary reaction 2(es2): } I3 + F \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} P + C$$

$$\text{Uncatalyzed reaction (uc): } S + F \underset{k_0}{\overset{k_0}{\rightleftarrows}} P + O$$

If the junction is absent, DNA reactants are identical to DNA products and ΔG^o_{OR} is zero. For elementary step 1, the substate **S** and intermediate ***I3*** share the same DNA hybridization domains **2:2*** and **3:3*** and ΔG^0 of the step 1, ΔG^o_{es1}, is determined by difference in binding energy of two toeholds, **1** and **4**. Because the step 2 is a reverse reaction of the step1, ΔG^0 of the step 2, ΔG^o_{es2}, is therefore determined by difference in binding energy of two toeholds, **4** and **1**. If the binding energy of the toehold **4** is lower than that of the toehold **1**, ΔG^o_{es2} is negative and $k_2 > k_{-2}$. Inversely, ΔG^o_{es1} is positive and $k_1 < k_{-1}$. Therefore, the step 1 is the rate limiting step, determining the rate of the overall reaction. Decreasing ΔG^o_{es1} can increase the ratio of k_1/k_{-1}, enhancing its kinetics and therefore the kinetics of the overall reaction. The introduction of a junction of into the substrate decreases ΔG^o_{es1}, thereby improving the kinetics of the overall catalytic reaction.

2.3.2 ΔG^0 of the overall catalytic reaction and elementary steps.

Table 2.1 Sequences of all the domains using in the TWJ mediated non-covalent DNA amplification.	
Domain	Sequence
D	CAT CAC CTC ACT CAA C CTA CCAC ATC TCC A-/3FAM/ 1 2 3 4
O-1'	/5IABkFQ/-TGG AGA T GTG GTA G GTT GAG TGA G 4* 3* 2*
O-1§	GCG AAG GAT GAG AT GTT GAG TGA G 5 2*
O-2	/5IABkFQ/-TGG AGA T GTG GTA G CTC ATC CTT CGC 4* 3* 5*
F	TGG AGA T GTG GTA G GTT GAG TGA G 4* 3* 2*
C	GTG GTA G GTT GAG TGA G GTG ATG 3* 2* 1*
§A 2 nt "TA" spacer is inserted between 2*and 5 on o-2 to relieve the strain in the junction	

Table 2.1 lists the sequences I designed for the substrate, fuel, and catalyst. Based on these sequences, I estimated ΔG^0 of formation of the substrate with and without junction, intermediate ***I3***, and product DNA (**Figure 2.2**). For the catalytic reaction using the substrate without the junction, the ΔG^o_{OR} is zero, and the ΔG^o_{es1} and ΔG^o_{es2} are 1.1 kcal/mole and -1.1 kcal/mole, respectively. By contrast, for the catalytic reaction using the substrate with the junction, the ΔG^o_{OR} is -7.5 kcal/mol, and the ΔG^0 values of the elementary steps 1 and 2 are -6.4 kcal/mole and -1.1 kcal/mole, respectively. Therefore, the overall reaction and two elementary steps of TWJ-mediated DNA catalytic reaction are energetically favorable, facilitating the kinetics of the catalytic reaction.

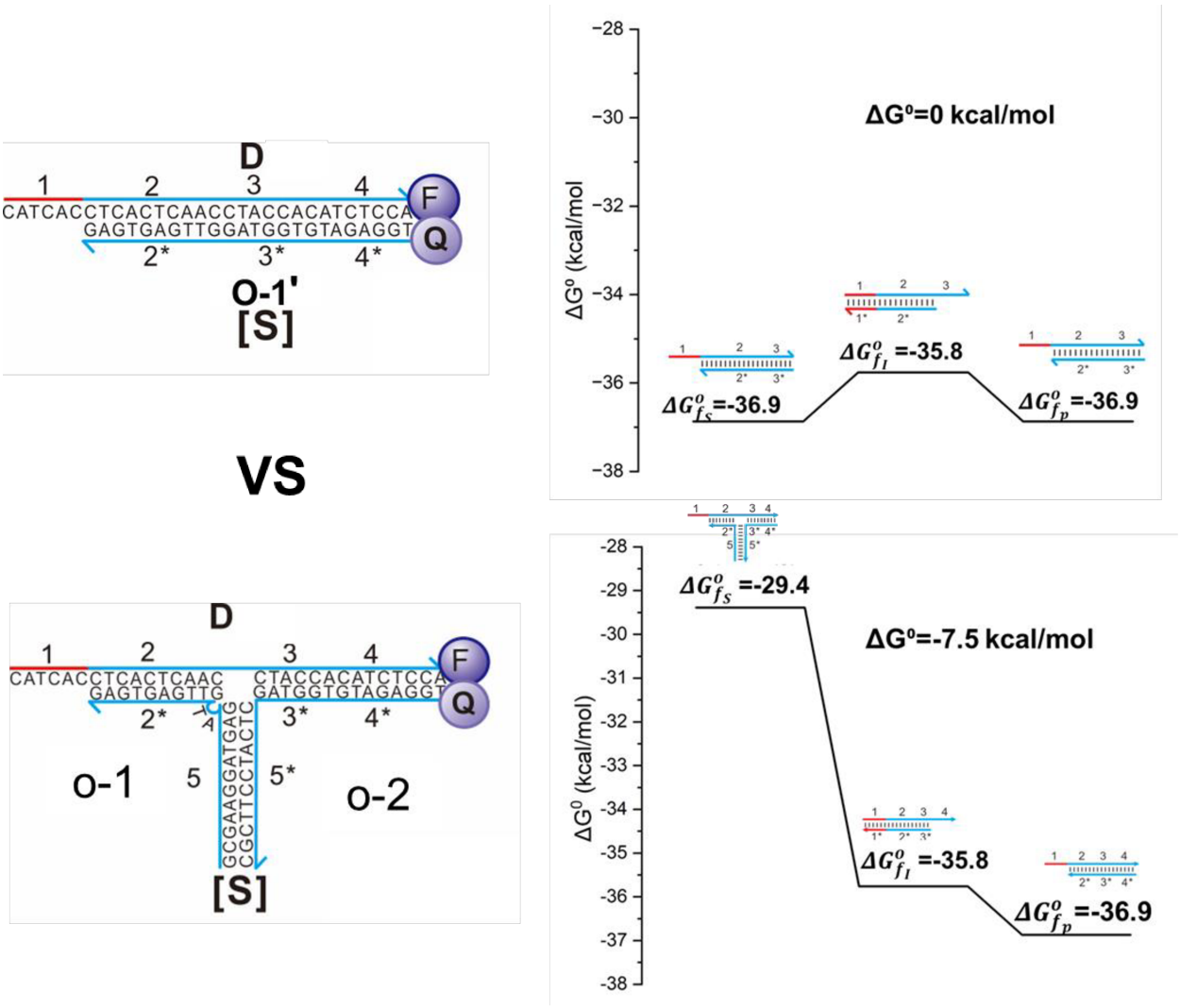


Figure 2.2 *$\Delta G^o_{f_S}$ with and without three-way junction and the free-energy changes of the elementary steps and the overall reaction.*

2.3.3 Reactions between the TWJ substrate and catalysts with different length

The ΔG^o_{es1} was estimated to be -6.4 kcal/mol, suggesting the forward reaction is favorable. This free energy (-6.4 kcal/mol) is approximately equivalent to the binding energy of four to five base pairs. If I truncate nucleotides from the 5'-end of the catalyst, the reaction between the TWJ substrate with truncated catalyst becomes less energetically favorable and the reaction equilibrium shifts towards the reverse reaction accordingly. To test this experimentally, I designed three

modified **C**: **C-2** has a one nucleotide extension while **C-3** and **C-4** have one and two nucleotides truncated at their 5'-ends as compared to the original catalyst DNA (**C**) (**Table 2.2**). I prepared four 20 μL reaction solutions each containing 300 nM the TWJ substrate and 300 nM **C**, **C-2**, **C-3**, or **C-4**. These reaction solutions were then incubated at room temperature for 30 min to allow the reactions to reach to equilibrium. The reaction solutions were then loaded onto 10% w/v PAGE gel to examine the generation of product DNA (**Figure 2.3**). The results of gel electrophoresis show that the TWJ substrate completely reacted with **C** and **C-2**, whereas some of the TWJ substrate remain unreacted for **C-3** and **C-4**. Additionally, the reaction with **C-4** left more amount of the TWJ substrate than with **C-3**. These results are consistent with the estimated ΔG^0 values of these reactions.

Table 2.2 Sequences of **C** and modified **C** with different reverse toehold length and their corresponding ΔG^o_{es1}.

Name	Sequence	ΔG^o_{es1} $\left(\Delta G^o_{f_{I3}} - \Delta G^o_{f_S}\right)$ (kcal/mol)
C-2	T GTG GTA GGT TGA GTG AGG TGA TG	-7.1
C-3	TG GTA GGT TGA GTG AGG TGA TG	-3.4
C-4	G GTA GGT TGA GTG AGG TGA TG	-3.1

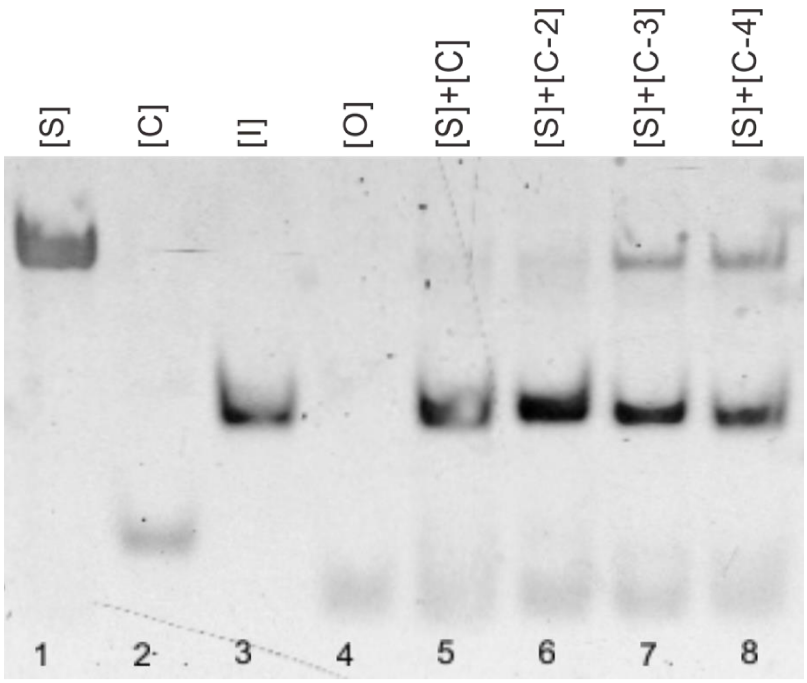


Figure 2.3 *PAGE image showing reactions between the TWJ substrate and catalysts with different length.* ***S*** *denotes substrate duplex;* ***C*** *denotes catalyst strand;* ***I*** *denotes intermediate duplex;* ***O*** *denotes output strand;* ***C-2, C-3****, and* ***C-4*** *are variants of the catalyst strand of different length.*

2.3.4 Typical amplification curves of TWJ-mediated DNA catalytic reaction

To determine whether the catalyst can catalyze the reaction of the TWJ substrate and fuel DNA, I prepared four reaction solutions each containing 25 nM TWJ substrate, 100 nM fuel DNA, and 0 pM, 50 pM, 100 pM, and 200 pM catalyst. **Figure 2.4** shows the amplification curves of these reaction solutions, where the fluorescence signal was converted to the amount of the product using a calibration curve established with a series of standard solutions containing varying concentrations of the product. Little fluorescence was observed when the catalyst was not added, suggesting the reaction rate of the TWJ substrate and fuel DNA is very slow in the absence of the catalyst. Steady fluorescence increases were observed over time for 50 pM, 100 pM, and 200 pM of the catalyst. The fluorescence increase is proportional to the concentration of the catalyst.

Collectively, these results indicate the catalyst can efficiently catalyze the reaction of the TWJ substrate and fuel, leading to amplified fluorescent signals.

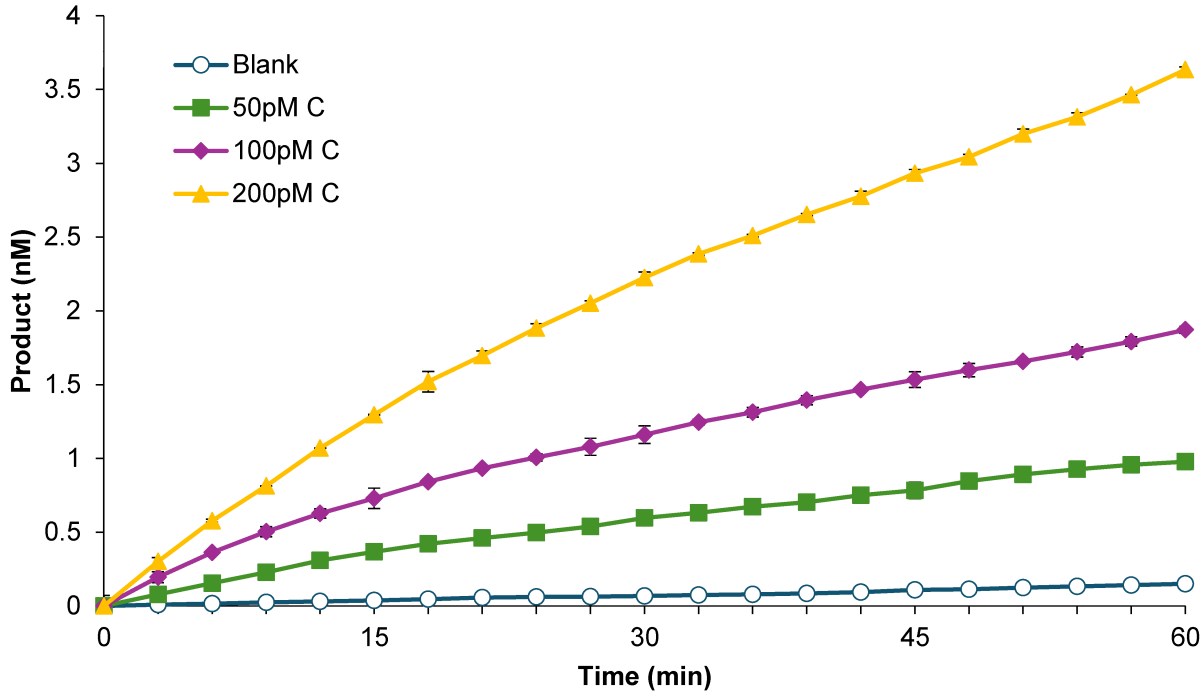


Figure 2.4 *Catalysis of three-way-junction mediated DNA catalytic reaction.*
The fluorescence signals was converted to nanomolar of the product using a calibration curve.

I further examined such catalytic reaction using PAGE (**Figure 2.5**). Three reaction solutions were prepared to contain 300 nM the TWJ substrate, 300 nM the fuel DNA, and 20 nM catalyst. After the solutions were incubated at room temperature for 10 min, 30 min, and 60 min, respectively, 20 μL of each solution was used for PAGE analysis. A blank was also prepared to contain all other components, except the catalyst. As the reaction time increased, the intensity of bands corresponding to the product DNA increased for three reaction solutions. Conversely, the intensity of bands corresponding to the TWJ substrate became fainter. No product DNA band was

observed for the blank. The results of PAGE analysis further confirmed that the catalyst can efficiently catalyze the reaction between the TWJ substate and the fuel.

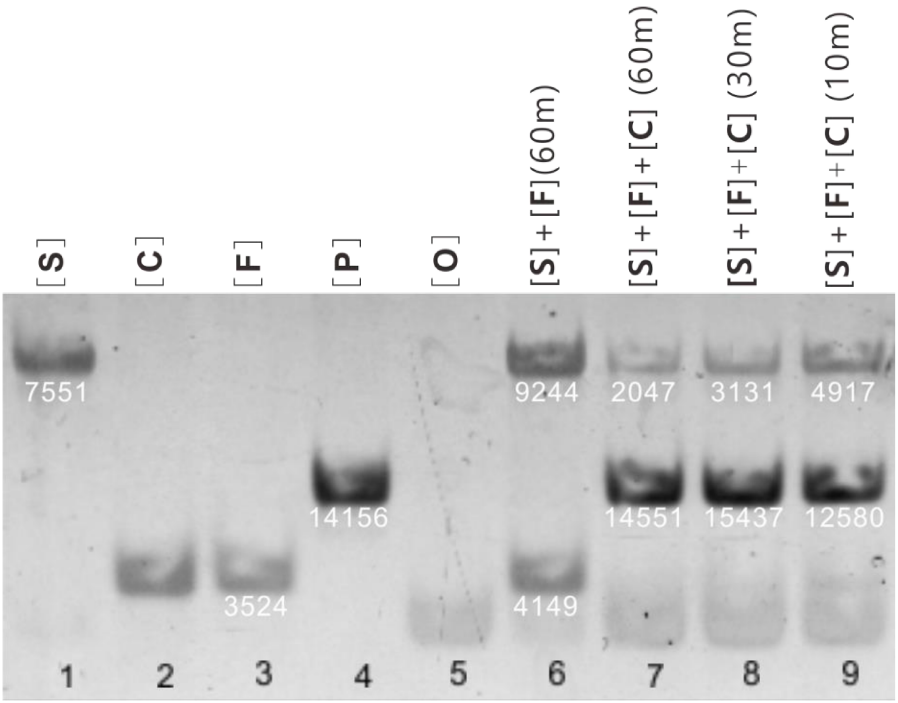


Figure 2.5 *PAGE image showing the three-way-junction mediated catalytic reaction.*
The band intensities, shown in white text, were determined using ImageJ. ***S*** *denotes the substrate;* ***C*** *denotes the catalyst;* ***F*** *denotes the fuel;* ***P*** *denotes the product;* ***O*** *denotes the output.*

2.3.5 Determining kinetics of TWJ mediated non-covalent DNA amplification.

I then determined the rate constants of elementary steps. Using the free energy changes ΔG^0 of the two elementary steps, I calculated k_1/k_{-1} and k_2/k_{-2} using the equation $\Delta G^0 = -RTln\ (k_f/k_r)$. The ratios of k_1/k_{-1} and k_2/k_{-2} were calculated to be 4.7×10^4 and 6.5, respectively. Therefore, for both elementary steps, k_f (k_1 or k_2) is significantly higher than k_r (k_{-1} or k_{-2}). To determine the k_f, I considered two elementary steps as irreversible reactions. The rate constants k_1 and k_2 were

determined to be 1.4×10^6 $M^{-1}s^{-1}$ and 1.3×10^6 $M^{-1}s^{-1}$, respectively. These values are comparable to those of the fastest toehold-mediated strand displacement reactions, indicating both elementary steps are fast, which benefits the amplification efficiency of the TWJ-mediated DNA catalytic reactions. Because k_1 is slightly higher than k_2, the elementary step 2 is the rate-limiting step of the catalytic reaction.

I also determined the rate constant k_0 of the uncatalyzed reaction, where the reaction solutions contained 25 nM TWJ substrate, and 100 nM fuel. The k_0 was determined to be 10 $M^{-1}s^{-1}$, suggesting that the reaction rate of the uncatalyzed reaction is very low. The uncatalyzed reaction is also referred as the leakage generating background. Therefore, the leakage of the TWJ-mediated DNA catalytic reaction is low. I found only about 2.3% of the TWJ substrate reacted with the fuel after an 8 h reaction. I then calculated k_2/k_0 to be 1.3×10^5, suggesting that the catalyst can improve the reaction kinetics by over five orders of magnitude.

.

2.3.6 Effect of junction location on TWJ mediated DNA catalytic reaction

I further investigated effect of the junction location on kinetics, leakage, and signal amplification of the TWJ mediated reaction. To facilitate the comparison and avoid the synthesis of multiple quencher-labeled **O-2** strands, I merged the **O-1** and **O-2** into a single strand **S-O** using a stem-loop structure (**Figure 2.6**). Consequently, the original TWJ substrate consisted of 3 DNA strands, **D**, **O-1**, and **O-2**, while the new TWJ substrate is comprised of only two DNA strands, **D** and **S-O**. The two types of the TWJ substrates share the same three-way junction structure and exhibited the same binding energy for the same junction location. To differentiate them, I named them as the 3S-TWJ substrate and 2S-TWJ substrate, respectively. Furthermore, I extended the **S-O** with a domain **5** (5'-AGAAGGTGAGTAGTG-3') at its 5'-end to facilitate the use of the same

reporter for detecting the release of all the output DNA **S-Os** (**Figure 2.6**). When 2S-TWJ substrate reacts with the catalyst, the **S-O** dissociates from **D** to exposes the domain **4***, enabling it to bind to the toehold on the DNA reporter. This initiates the toehold-mediated strand displacement reaction, generating the fluorescence signal.

I designed ten **S-O** strands with different junction locations (**Figure 2.6, Table2.4**). Each **S-O** is denoted as **S-O (m, n)**, indicating there are m nucleotides between the junction and the 3'-end of domain **2*** and n nucleotides between the junction and the 5'-end of domain **4***. Therefore, each **S-O** strand contains a 5' arm and a 3' arm, as shown in **Table 2.3.** Subsequently, I assessed the free energy, ΔG^o_{fs}, of each 2S-TWJ substrate and binding energy of its two arms. Notably, our estimation revealed that the junction location only slightly impacts the ΔG^o_{fs}, as the difference between the lowest ΔG^o_{fs} (-32.8 kcal/mol) and the highest ΔG^o_{fs} (-31.2 kcal/mol) is only 1.6 kcal/mol. In contrast, the binding energy of arms exhibits substantial variation across different junction locations.

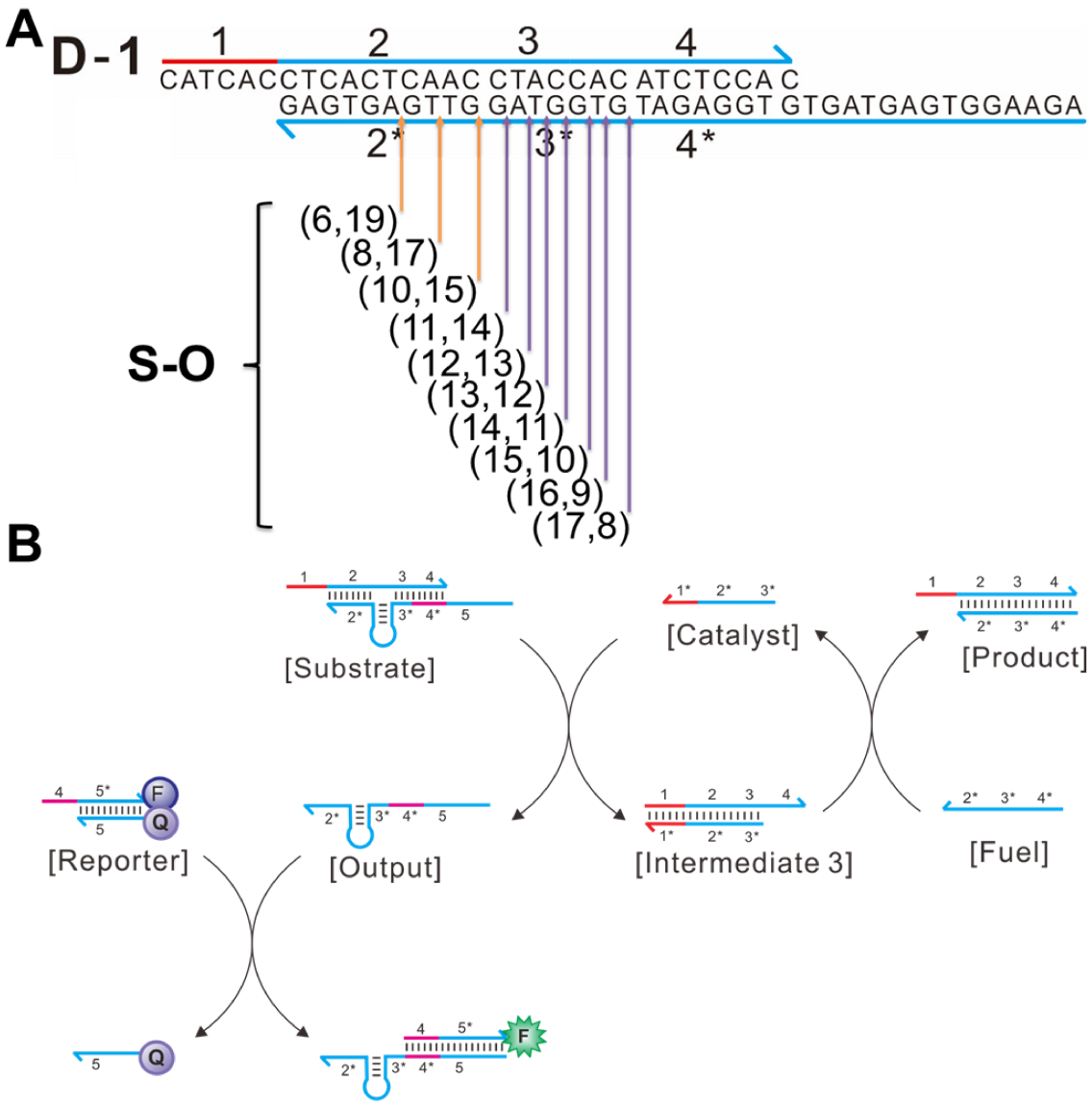


Figure 2.6 *Schemes showing the design of 2S-TWJ substrates with varying junction locations and DNA catalytic reaction using the 2S-TWJ substrate.*

(A) Each S-O is denoted as S-O (m, n), where m denotes the number of nucleotides between the junction and the 3'-end of domain ***2**** *and n refers to the number of nucleotides between the junction and the 5'-end of domain* ***4****. (B) The proposed reaction pathway of the TWJ-mediated DNA catalytic reaction using the 2S-TWJ substrate.*

Table 2.3 *Sequence o used for formation of 2S-TWJ substrates with varying junction locations and their free energies and binding energies of the two arms*

	Sequence	ΔG_f^o of m arm (kcal/mol)	ΔG_f^o of n arm (kcal/mol)	ΔG_{fs}^o (kcal/mol)
D-1	CAT CAC CTC ACT CAA CCT ACC ACA TCT CCA C	N/A		
S-O (6,19)	AGA AGG TGA GTA GT G TGGAGAT G TGG TAG GTT G **CTCAT** AAAAA **ATGAG** TA AGTGAG	-29.2	-9.2	-31.8
S-O (8,17)	AGAAGGTGAGTAGT G TGGAGAT GTGGTAG GT **CTCAT** AAAAA **ATGAG** TA TGAGTGAG	-26.0	-12.5	-32.2
S-O (9,16)	AGAAGGTGAGTAGT G TGGAGAT GTGGTAG G **CTCAT** AAAAA **ATGAG** AT TTGAGTGAG	-25.2	-12.6	-31.2
S-O (10,15)	AGAAGGTGAGTAGT G TGGAGAT GTGGTAG **CTCAT** AAAAA **ATGAG** AT GTTGAGTGAG	-22.7	-15.2	-31.7
S-O (11,14)	AGAAGGTGAGTAGT G TGGAGAT GTGGTA **CTCAT** AAAAA **ATGAG** AT G GTTGAGTGAG	-21.1	-18.2	-32.8
S-O (12,13)	AGAAGGTGAGTAGT G TGGAGAT GTGGT **CTCAT** AAAAA **ATGAG** TA AG GTTGAGTGAG	-20.6	-19.1	-32.5
S-O (13,12)	AGAAGGTGAGTAGT G TGGAGAT GTGG **CTCAT** AAAAA **ATGAG** AT TAG GTTGAGTGAG	-19.3	-19.5	-32.1
S-O (14,11)	AGAAGGTGAGTAGT G TGGAGAT GTG **CTCAT** AAAAA **ATGAG** AT GTAG GTTGAGTGAG	-16.8	-21.0	-31.7
S-O (15,10)	AGAAGGTGAGTAGT G TGGAGAT GT AA **CTCAT** AAAAA **ATGAG** GGTAG GTTGAGTGAG	-14.7	-23.6	-31.9
S-O (16,9)	AGAAGGTGAGTAGT G TGGAGAT G **CTCAT** AAAAA **ATGAG** AA TGGTAG GTTGAGTGAG	-14.0	-24.4	-31.7
S-O (17,8)	AGAAGGTGAGTAGT G TGGAGAT AA **CTCAT** AAAAA **ATGAG** GTGGTAG GTTGAGTGAG	-11.9	-16.9	-31.7
Sig-Q-1	5IABkFQ/ -AGA AGG TGA GTA GTG	N/A		
Sig-F-1	ATC TCC A CAC TAC TCA CCT TCT-/FAM3	N/A		

Each S-O is denoted as S-O (m, n), where m denote the number of nucleotides between the junction and the 3'-end of domain 2* and n the number of nucleotides between the junction and the 5'-end of domain 4*.

Domain are separated by underline: domain 5, m arm, stem loop and spacer, n arm

I then determined the rate constants k_1 and k_0 for the reactions using these ten 2S-TWJ substrates with different junction locations (**Figure 2.7**). The DNA catalytic reactions employing these 2S-TWJ substrates have a same k_2 value due to their identical elementary step 2. Notably, the impact of the junction location on k_0 is stronger than on k_1. Junctions proximal to the 3'-end of domain **2*** and 5'-end of domain **4*** resulted in larger k_0 values, whereas k_0 decreased as the junction moves towards the middle location of the DNA hybrid between domains **2***, **3***, and **4*** with their complements. The 2S-TWJ substrate with **S-O** (14, 11) exhibited the lowest k_0 values (3.4 M^{-1} s^{-1}), which is approximately 50-fold lower than the highest k_0 (178 M^{-1} s^{-1}) the 2S-TWJ substrate with **S-O** (9, 16).

The effect of the junction location on k_1 is relatively lower compared to k_0. The highest k_1 is approximately three times higher than the lowest. Except for the junction (6, 19), which led to the highest leakage, junctions near the reverse toehold **4** generally exhibit higher k_1 values than those near the forward toehold **1**. This suggests that junctions near the forward toehold are less favorable for the kinetics of the elementary step 1. The best k_1/k_0 was obtained for the junctions (14, 11) and (15, 10). Additionally, I determined the amplification efficiency of the substrates with different junction locations (**Figure 2.8**), revealing a similar trend to that observed for k_1.

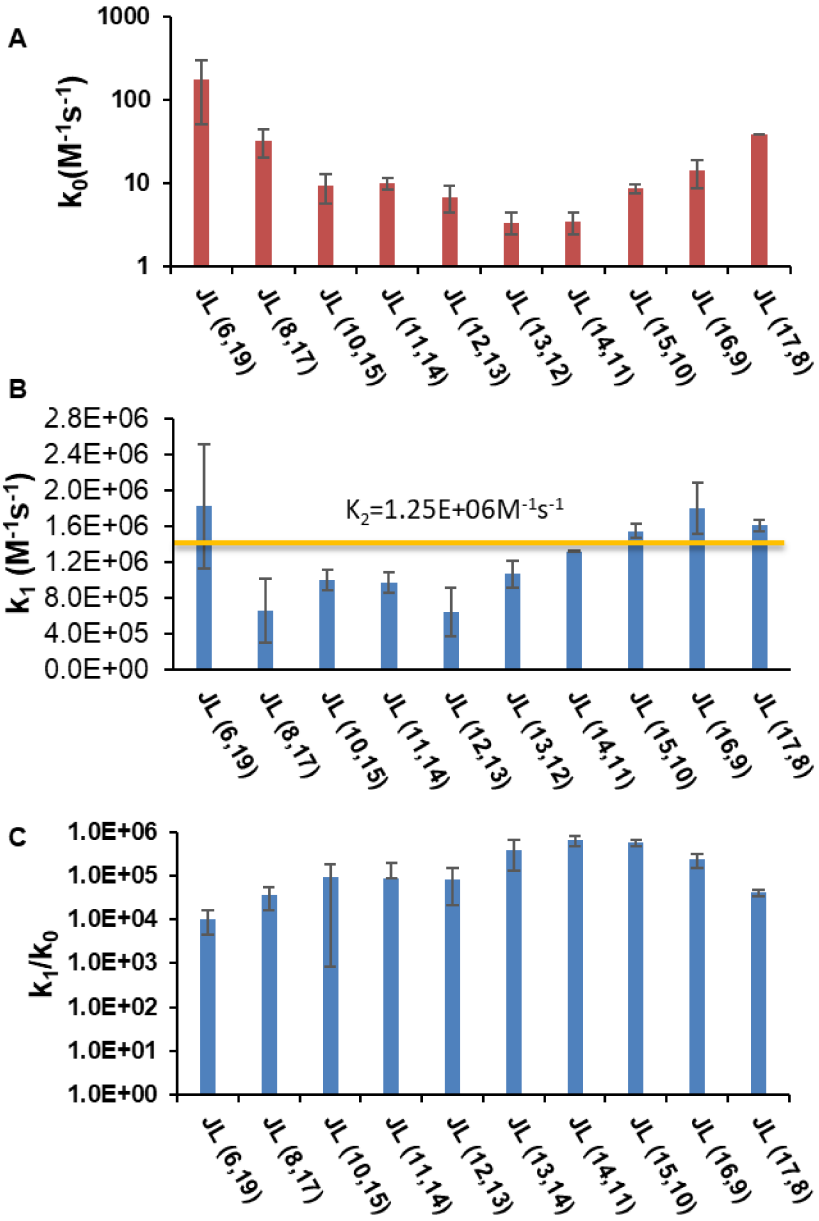


Figure 2.7 *Effect of the junction location on reaction kinetics.*

*(**A**) k_0 of each substrate (**B**) k_1 of each substrate. (**C**) k_1/k_0 of different junction locations. **JL** (m,n) denotes the junction locations, where m denotes the number of nucleotides between the junction and the 3'-end of domain **2*** and n refers to the number of nucleotides between the junction and the 5'-end of domain **4***.*

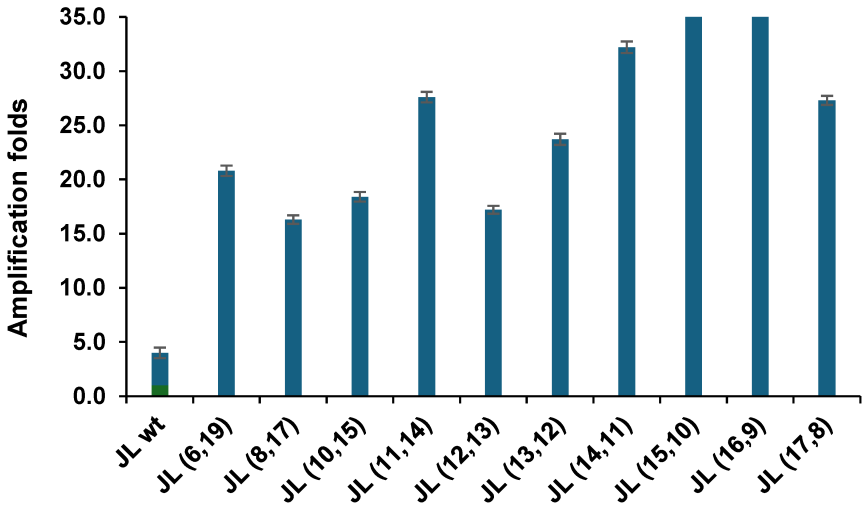


Figure 2.8 *Effect of the junction location on amplification efficiency of the TWJ mediate DNA catalytic reaction.* ***JL*** *(m,n) denotes the junction locations, where m denotes the number of nucleotides between the junction and the 3'-end of domain* ***2**** *and n refers to the number of nucleotides between the junction and the 5'-end of domain* ***4****.*

2.3.7 Comparing TWJ substrate and non-TWJ Substrate for DNA catalysis

To determine the improvement in amplification of the TWJ-mediated DNA catalytic reaction compared to the reaction using a substrate without TWJ, I compared the performance of two reactions catalyzed by 200 pM of the catalyst (**Figure 2.9**). Both reactions were conducted under the same conditions: 25 nM of TWJ substrate or substrate without TWJ, 100 nM fuel, and 100 pM catalysts in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween. The reactions were monitored with real-time fluorescence detection. The fluorescent signals were converted to molar amounts of the product using a calibration curve. The TWJ-mediated DNA catalytic reaction resulted in approximately 7 nM product, representing 35-fold signal amplification in a one-hour reaction. In contrast, the DNA catalytic reaction using the substrate without TWJ produced about 0.5 nM product, equivalent to 2.5-fold signal amplification.

Therefore, compared to the DNA catalytic reaction using the substrate without TWJ, the TWJ-mediated DNA catalytic reaction improved the amplification efficiency by 14-folds.

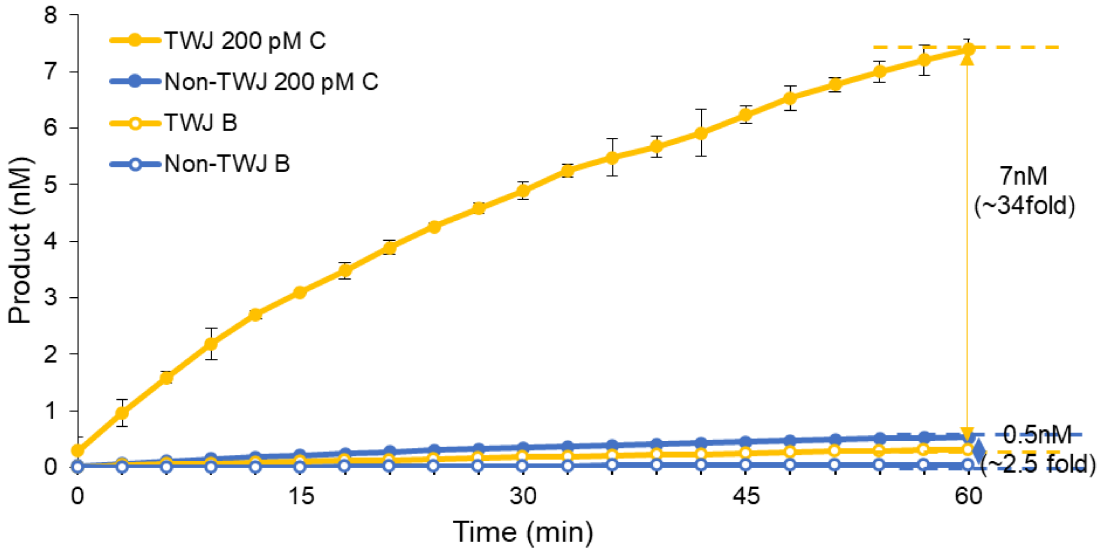


Figure 2.9 *Comparison of the TWJ substrates with the non-TWJ substrates for DNA catalysis. The catalyst is denoted as* ***C*** *while B denotes the blank.*

2.3.8 Determining specificity of TWJ mediated DNA catalytic reaction.

I further determined the specificity of the TWJ-mediated DNA catalytic reaction. I designed five single-mismatch variants of the catalyst, each with a mismatch base at a different site (**Table 2.4**). I then evaluated the performance of the TWJ-mediated DNA catalytic reaction catalyzed by these five variants as well as the original catalyst (**Figure 2.10**). Discrimination factor(DF) was calculated using the formula DF = $\frac{(Signal_C - Background)/[C]}{(Signal_{varient} - Background)/[Variant]}$. The specificity increased as the mismatch base was positioned further from the junction. The largest discrimination factor was (DF=325) was observed when the mismatch base was in the toehold binding domain **1***. The

presence of mismatch in **1*** disrupts its hybridization to the toehold **1**, raising the free-energy change of toehold hybridization, reducing dissociation half-life and the likelihood of initiating TE.

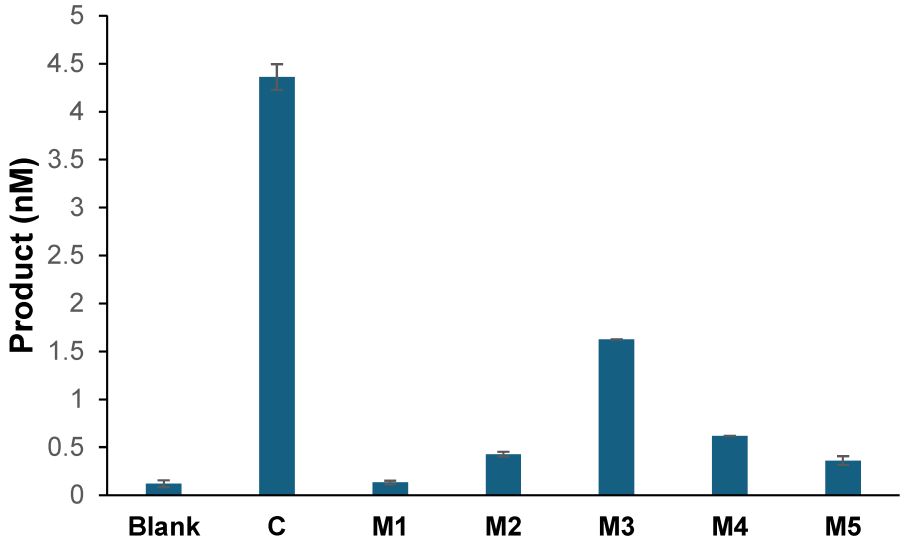


Figure 2.10 *Products generated by the catalyst,* ***C****, and five single-mismatch variants,* ***M1-M5****.*

Table 2.4 *Sequences of single mismatch variants and their corresponding discrimination factors*		
Name	Sequence	Discrimination factor (DF)
M1	GTG GTA GGT TGA GTG AGG T**T**A TG	325
M2	GTG GTA GGT TGA GTG **T**GG TGA TG	13.7
M3	GTG GTA GGT TG**T** GTG AGG TGA TG	2.8
M4	GTG GTA G**T**T TGA GTG AGG TGA TG	8.5
M5	GTG **T**TA GGT TGA GTG AGG TGA TG	17.5

2.3.9 Using TWJ-mediated DNA catalytic reaction as an assay for amplified detection of the catalyst

Because the catalyst facilitates the reaction between the TWJ substrate and fuel, generating an amplified fluorescence signal, the TWJ-mediated DNA catalytic reaction can be directly used as an assay for the amplified detection of the catalyst. To achieve high detection sensitivity, I first studied the effect of the substrate and fuel on the amplification efficiency and leakage of the catalytic reaction (**Figure 2.11, 2.12**), using 100 pM of the catalyst. Increasing the substrate concentration from 25 nM to 100 nM led to a continuous increase in signal amplification but resulted in a large increase in background. The 25 nM substrate concentration provided the best signal-to-background ratio (**Figure 2.11B**). When the fuel concentration was increased from 25 nM to 200 nM, the signal amplification was continuously enhanced, while the background was only slightly affected. The 100 nM and 200 nM fuels resulted in the highest and comparable signal to background ratios (**Figure 2.12B**). Therefore, the substrate is the main source of leakage. This is because the leakage is often attributed to error in synthesis of the DNA oligonucleotide, which results in malformation of the substrates and leakage.

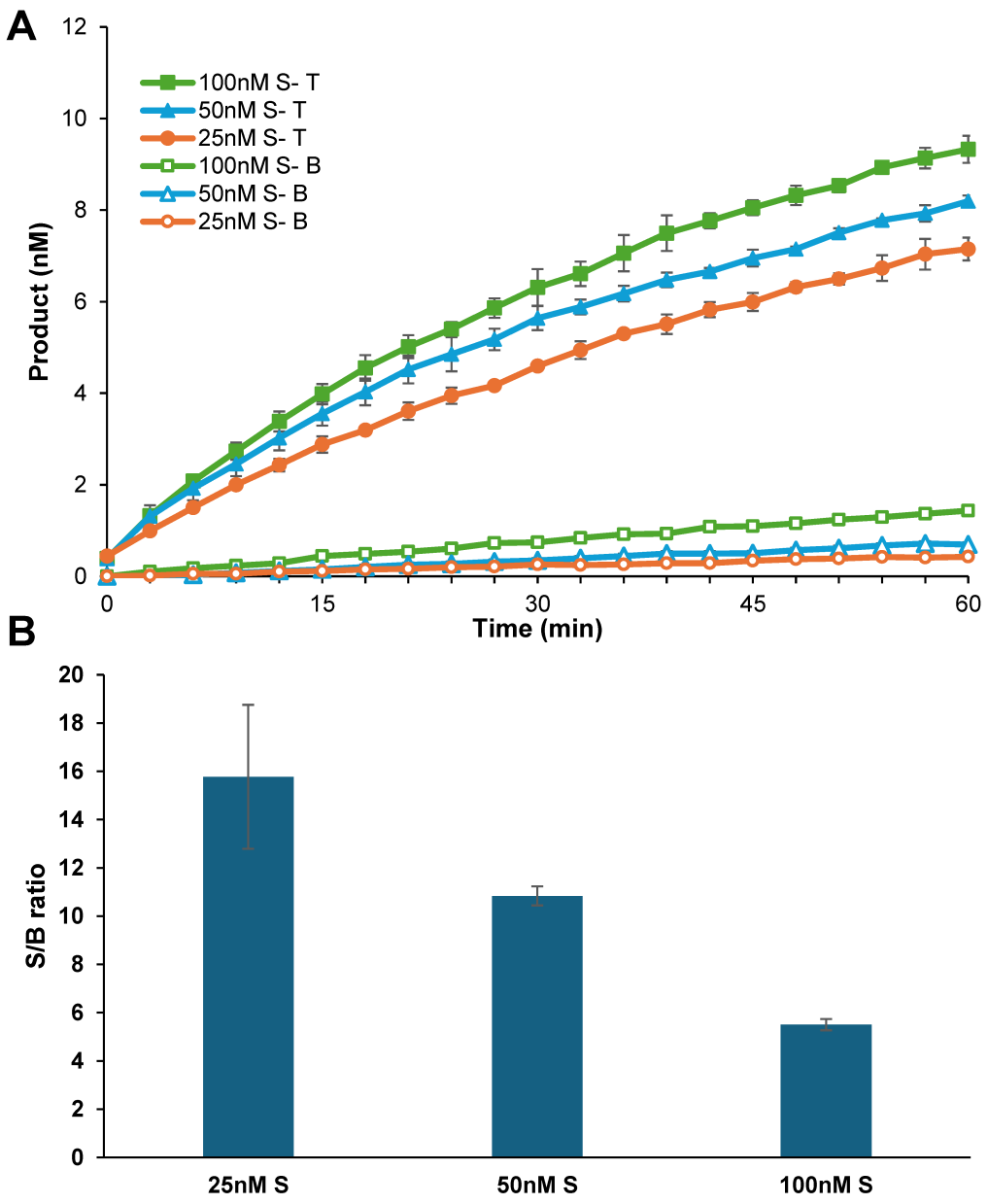


Figure 2.11 *Impact of the substrate concentration on the signal amplification.*

(A) Amplification curves, ***S*** *denotes the substrate,* ***T*** *denote reactions where 100pM of the catalyst was added, and* ***B*** *denotes the blank. (B) Signal to background ratio.*

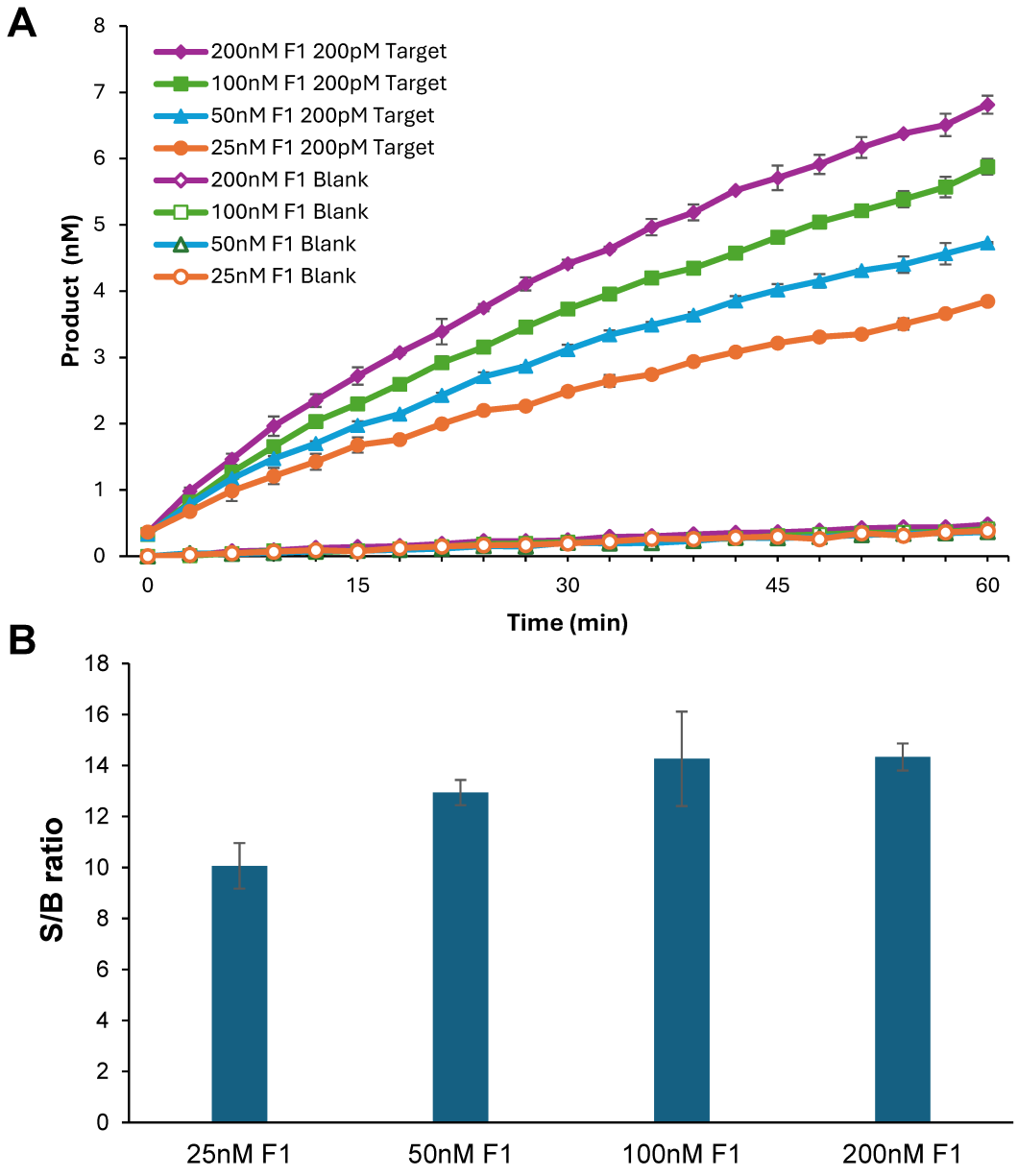


Figure 2.12 *Impact of the fuel concentration on the signal amplification.*

(A) Amplification curves, ***F1*** *denotes the fuel. (B) Signal to background ratio.*

I chose 25 nM substrate and 100 nM fuel to determine the dynamic range and limit of detection of the catalytic reaction as an assay. A linear dynamic range from 5 pM to 500 pM the catalyst was obtained (**Figure 2.13**). The assay was able to detect as little as 2 pM of the catalyst

within 1 hour at room temperature without the need for any enzymes. This makes the assay promising for point-of-care and on-site applications as well as in situ signal amplification.

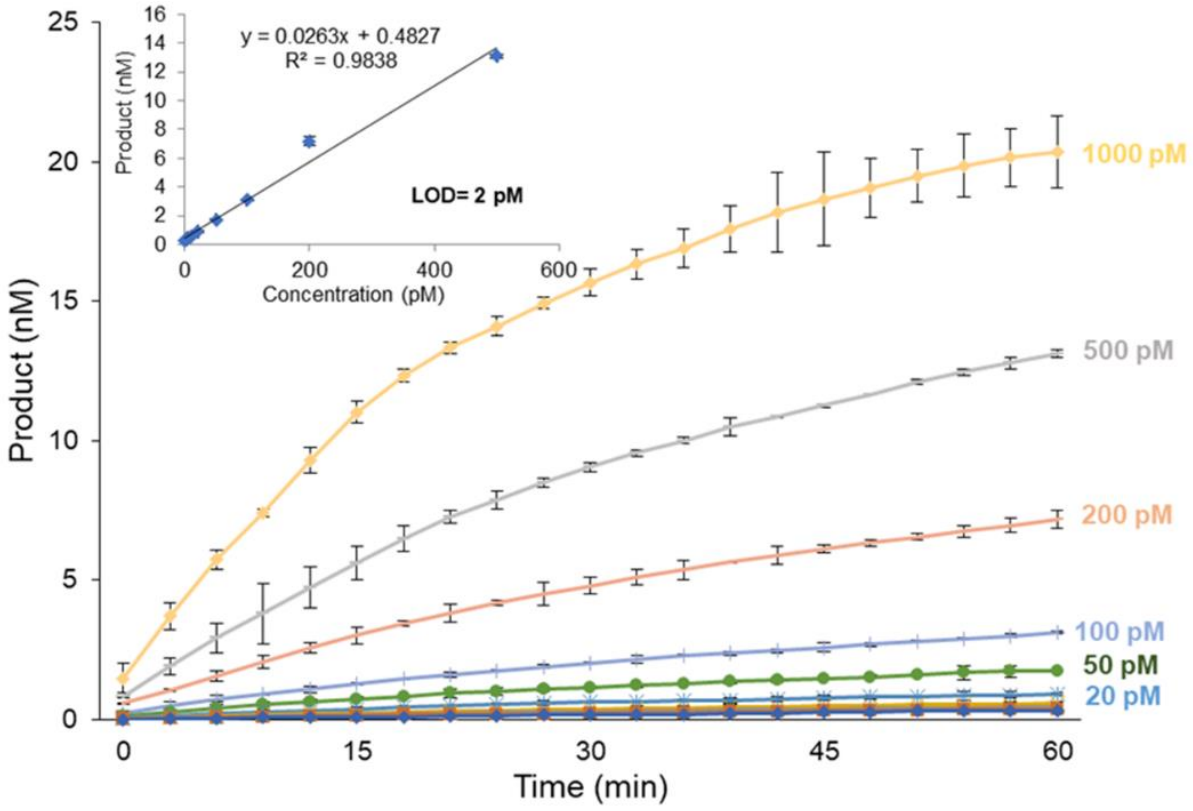


Figure 2.13 *Dynamic range and limit of detection TWJ mediated DNA catalytic reaction.*

2.3.10 Construction of two-layer TWJ mediated DNA catalytic reaction.

To demonstrate the capability for constructing multi-layered reaction cascades and improve amplification efficiency, I developed a two-layer TWJ-mediated DNA catalytic reaction system where two TWJ-mediated DNA catalytic reactions were paired (**Table 2.5, Figure 2.14**). The key design principle involves using a catalyst to initiate the first TWJ-mediated DNA catalytic reaction, which generates a layer 1 output DNA that serves as the catalyst for the second TWJ-mediated DNA catalytic reaction. Consequently, one catalyst molecule can produce multiple layer

1 output DNA strands through the layer 1 DNA catalytic reaction, each of which further generates multiple layer 2 output DNA strands through the layer 2 DNA catalytic reaction, achieving tandem amplification. Specifically, the catalyst catalyzes the reaction of the layer 1 substrate with the layer 1 fuel, generating the layer 1 output DNA. The layer 1 output DNA is a hybrid of layer 1 **O-1** and layer 1 **O-2**, which cannot directly serve as the catalyst for the layer 2 catalytic reaction. Therefore, I designed a single-stranded translator that reacts with layer 1 **O-2**, releasing layer 1 **O-1**, which is used as the catalyst for the layer 2 catalytic reaction. The layer 1 **O-1** then catalyzes the reaction between the layer 2 substrate and layer 2 fuel, generating layer 2 output DNA and layer 2 product. Because a pair of fluorophore and quencher are labeled at the blunt end of the layer 2 substrate. One cycle of the catalyzed layer 2 reaction lights up a fluorophore molecule. Thousands of fluorophore molecules can be turned on in response to a single catalyst molecule.

Table 2.5 *Sequences of DNA oligonucleotides used in the two-layer amplification.*

Domain	Sequence
D-2	CAT CAC CTC ACT CAA C CTA CCA C ATC TCC A 1 2 3 4
Translator	GCG AAG TAG GAG CTA CCA 5 3
Layer 2 D	CAC TCA ACT A CTC ATC CTT CGC ATC ACC T -/3FAM/ 2a 2b 5b* 5a* 6
Layer 2 O-1	CCG TCT TGT AGC AT GAT GAG TAG T 7 5b 2b*
Layer 2 O-2	5IABkFQ/-AGG TGA T GCG AAG GCT ACA AGA CGG 6* 5a 7*
Layer 2 Fuel	AGG TGA T GCG AAG GAT GAG TAG T 6* 5a 5b 2b*

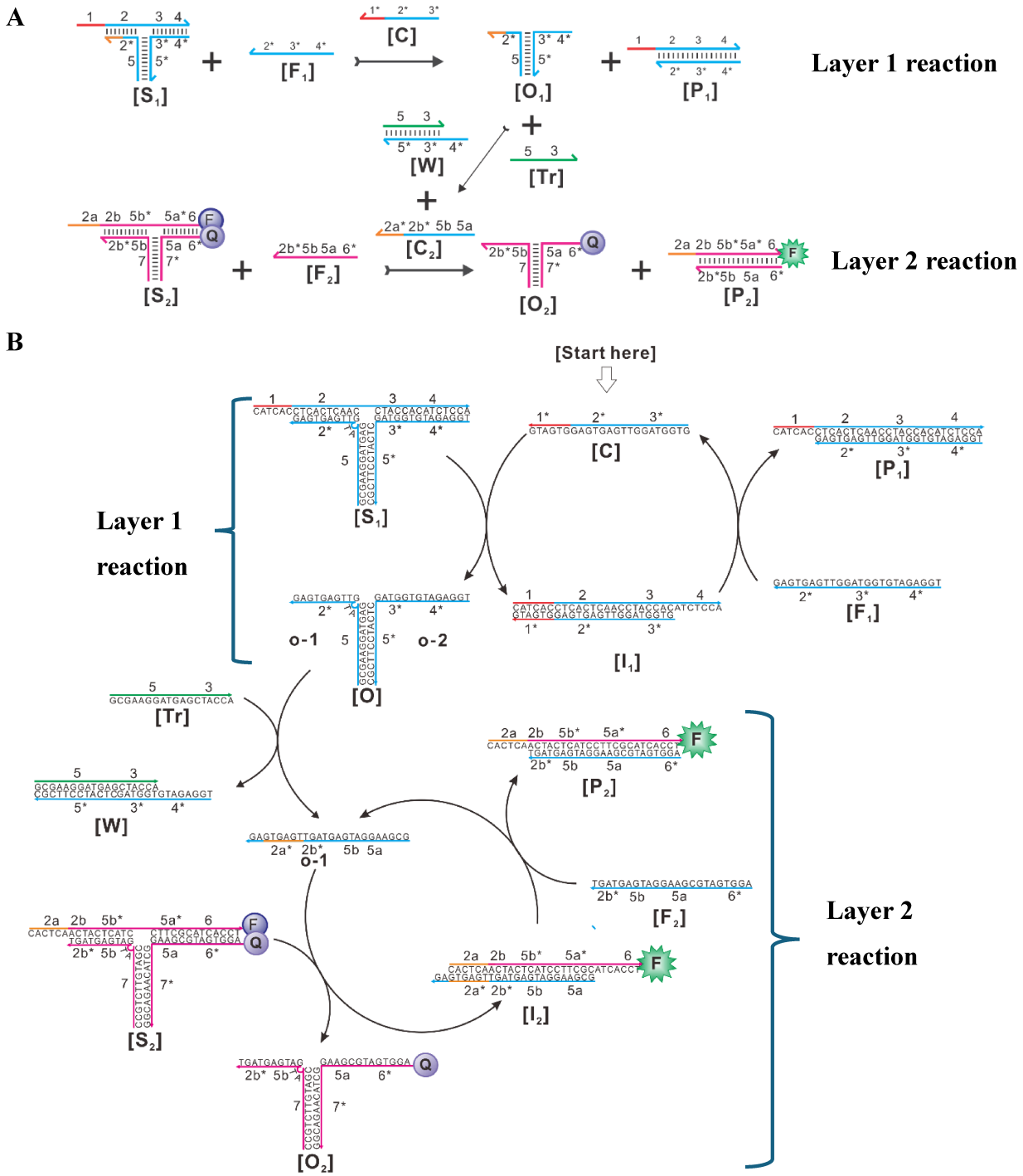


Figure 2.14 *Scheme of the two-layer TWJ mediated DNA catalytic reaction.*

(A) The overall two layered catalytic reaction. (B) The detailed reaction pathway of the two-layered catalytic reaction. ***S*** *denotes the substrate;* ***C*** *denotes the catalyst;* ***O*** *denotes the output;* ***I*** *denotes the intermediate;* ***P*** *denotes the product;* ***Tr*** *denotes the translator strand; and* ***W*** *denotes the waste.*

I first examined the amplification efficiency of the two-layer catalytic reaction, using 10 pM catalyst and conducting the reaction for three hours (**Figure 2.15**). A 250-fold signal amplification was observed after one hour, 1000-fold after two hours, and 1950-fold after three hours. I also compared the one-layer catalytic reaction with the two-layer catalytic reaction for signal amplification using 200 pM catalysts (**Figure 2.16**). The two-layer catalytic reaction achieved 235-fold signal amplification, while the one-layer catalytic reaction only achieved 25-fold amplification. Therefore, the two-layer catalytic reaction significantly improved the amplification efficiency compared to the one-layer reaction.

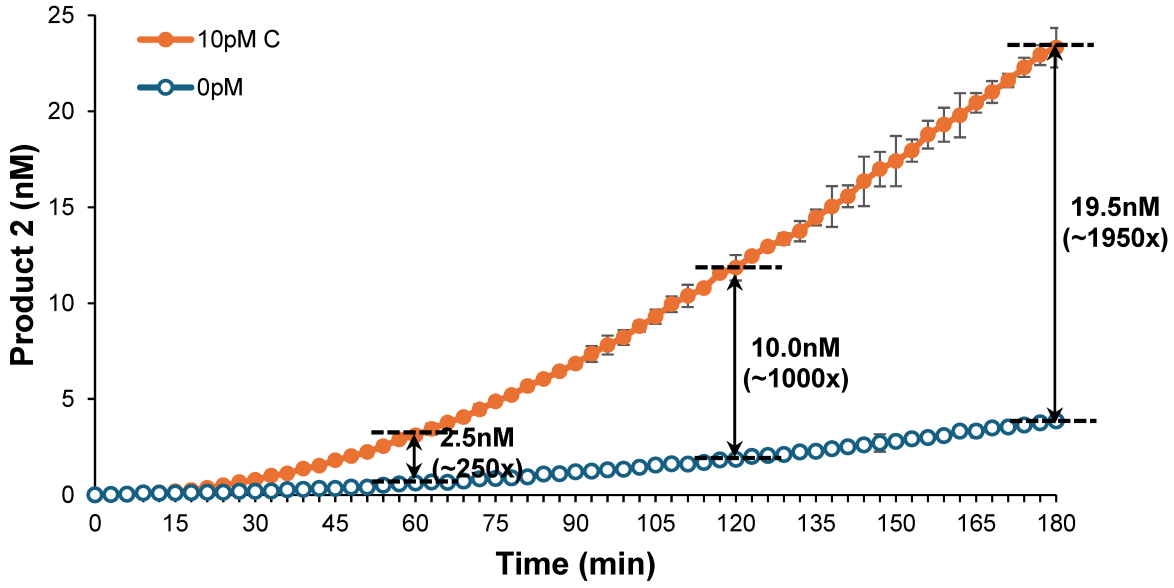


Figure 2.15 *Amplification efficiency of the two-layer TWJ mediated DNA catalytic reaction.* ***C*** *denotes the catalyst.*

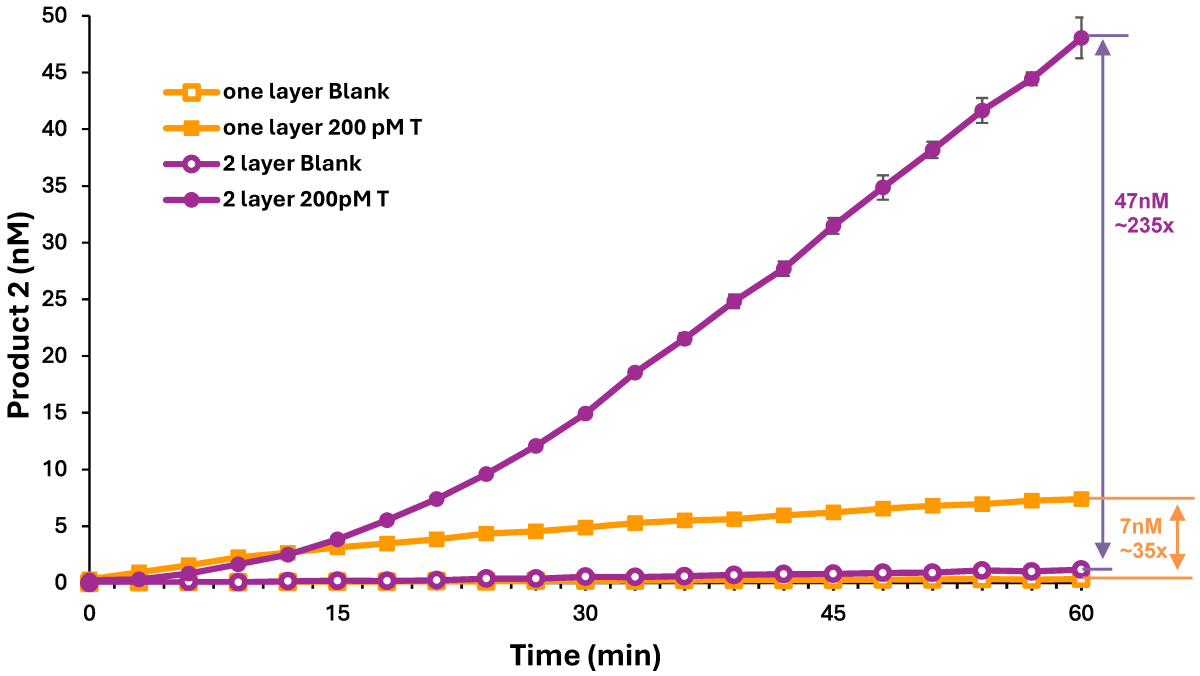


Figure 2.16 *Comparison of the amplification efficiency of the two-layer and one-layer catalytic reaction.* ***T*** *denotes target or the catalyst* ***C****.*

I then used the two-layer catalytic reaction to detect varying concentrations of the catalyst (**Figure 2.17**). A dynamic range between 0.5 pM and 100 pM was obtained, and the reaction was able to detect as low as 0.3 pM catalyst.

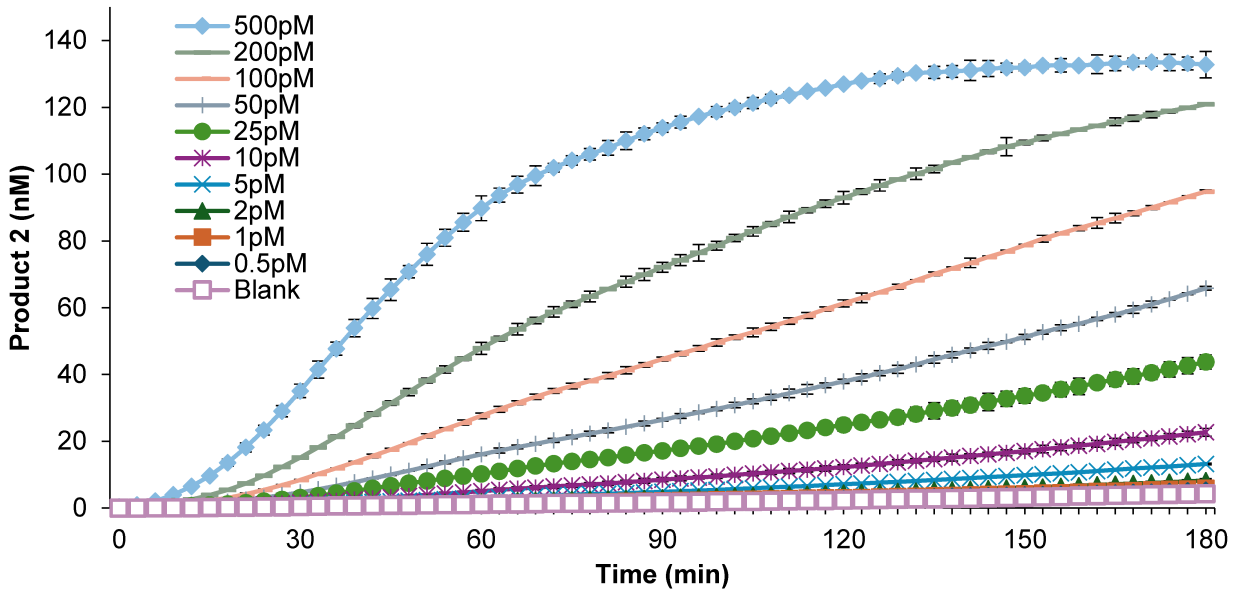


Figure 2.17 *Amplification curves of the two-layer catalytic reaction for detection of varying concentrations of the catalyst.*

2.4 Conclusion

I developed a simple and fast DNA catalytic reaction, termed the TWJ-mediated catalytic reaction, by introducing a TWJ into the substrate. This reaction overcomes the low amplification efficiency of the toehold exchange catalytic reaction, which produces two products identical to the substrate and fuel. Consequently, the free energy change of the toehold exchange catalytic reaction is zero, unfavorable for reaction kinetics. Introducing a TWJ into the substrate destabilizes it by disrupting normal base-stacking in dsDNA, thereby increasing the free energy of substrate. As the product DNA remains the same when the TWJ substrate is used, the free energy change of the overall reaction decreases to about -7 kcal/mol, making the reaction spontaneous.

I have demonstrated that the TWJ-mediated DNA catalytic reaction exhibits an amplification efficiency 14 times higher than the toehold exchange catalytic reaction, suggesting that increasing the free energy of the substrate can enhance the kinetics of catalytic reactions.

Another advantage of the TWJ structure is that it does not alter the base pairs of the original dsDNA, ensuring that no nucleotides in toehold-containing DNA are unpaired, which reduces the background leakage between the substrate and fuel. I found that when the binding of the two arms is stable, the background leakage remains low.

Due to its high amplification efficiency and low leakage, the TWJ-mediated catalytic reaction can detect as low as 2 pM catalyst at room temperature without the need for any enzymes. Additionally, I constructed a two-layer TWJ-mediated catalytic reaction to further improve amplification efficiency. This two-layer system exhibited a 1950-fold signal amplification in a three-hour reaction, lowering the limit of detection to 0.3 pM for the catalyst.

In addition to its high amplification efficiency, the TWJ substrate is easy to prepare through the simple annealing of two or three strands without requiring any purification. The high sensitivity, enzyme-free nature, and ease of use make the TWJ-mediated DNA catalytic reaction promising for point-of-care and on-site applications as well as *in situ* signal amplification.

2.5 Reference

1. Peng, H., Newbigging, A. M., Reid, M. S., Uppal, J. S., Xu, J., Zhang, H., et al. (2020). Signal amplification in living cells: A review of microRNA detection and imaging. Anal Chem. 92, 292.
2. Zhang, D. Y., & Winfree, E. (2010). Robustness and modularity properties of a non-covalent DNA catalytic reaction. Nucleic Acids Res. 38(12), 4182.
3. Simmel, F. C., Yurke, B., & Singh, H. R. (2019). Principles and applications of nucleic acid strand displacement reactions. Chem Rev. 119(8), 6326.
4. Turberfield, A. J., Mitchell, J. C., Yurke, B., Mills, A. P., Blakey, M. I., & Simmel, F. C. (2003). DNA fuel for free-running nanomachines. Phys Rev Lett. 90(11), 118102.
5. Seelig, G., Yurke, B., & Winfree, E. (2006). Catalyzed relaxation of a metastable DNA fuel. J Am Chem Soc. 128(30), 12211.

6. Dirks, R. M., Pierce, N. A. (2004). Triggered amplification by hybridization chain reaction. Proc Natl Acad Sci USA. 101, 15275.
7. Yin, P., Choi, H. M. T., Calvert, C. R., Pierce, N. A. (2008). Programming biomolecular self-assembly pathways. Nature. 451, 318.
8. Venkataraman, S., Dirks, R. M., Rothemund, P. W. K., Winfree, E., & Pierce, N. A. (2007). An autonomous polymerization motor powered by DNA hybridization. Nat Nanotechnol. 2(8), 490.
9. Zhang, D. Y., Turberfield, A. J., Yurke, B., Winfree, E. (2007). Engineering entropy-driven reactions and networks catalyzed by DNA. Science. 318, 1121.
10. Zhang, D. Y., Winfree, E. (2009). Control of DNA strand displacement kinetics using toehold exchange. J Am Chem Soc. 131, 17303.
11. Li, B., Ellington, A. D., & Chen, X. (2011). Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. Nucleic Acids Res. 39(15), e110
12. Leontis, N. B., Kwok, W., Newman, J. S. (1991). Stability and structure of three-way DNA junctions containing unpaired nucleotides. Nucleic Acids Res. 19, 759.
13. Carr, C. E., Marky, L. A. (2017). Investigation of the melting behavior of DNA three-way junctions in the closed and open states. Biophys J. 113, 529.
14. SantaLucia Jr., J, Hicks D. (2004). The thermodynamics of DNA structural motifs. Annu Rev Biophys. 33, 415.
15. Fornace, M. E., Huang, J. C., Newman, T., Porubsky, N. J., Pierce, M. B., Pierce, N. A. (2022). NUPACK: analysis and design of nucleic acid structures, devices, and systems. ChemRxiv.
16. Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., Pierce, N. A. (2011). NUPACK: analysis and design of nucleic acid systems. J Comput Chem. 32, 170.
17. Schneider, C., Rasband, W. & Eliceiri, K. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 9, 671.

3. Chapter 3: Mismatch-aided DNA catalytic reaction

3.1 Introduction:

Through the development of the TWJ-mediated DNA catalytic reaction, I have demonstrated that increasing the free energy of the substrate can make the overall catalytic reaction energetically favorable, thereby improving the kinetics and amplification efficiency of the catalytic reaction. However, the TWJ structure does not allow for significant alteration of the free energy of the substrate[1,2]. For instance, I designed ten TWJ substrates with varying junction locations, but the difference in free energy between the lowest (-32.8 kcal/mol) and the highest (-31.2 kcal/mol) is only 1.6 kcal/mol. Therefore, the use of TWJ substrates does not permit a systematic study of the relationship between the free energy of the substrate and the kinetics of both catalyzed and uncatalyzed reactions. A new approach is needed to effectively modulate the free energy of the substrate.

I chose to introduce mismatches into the substrates to modulate their free energy for two main reasons[3, 4]. First, mismatches are very flexible in altering the free energy of dsDNA, and the free energy of the substrate can be effectively tuned through changing of the mismatched bases and mismatch number. Previous studies have shown that the impact of DNA mismatches on the free energy of dsDNA follows the order of C-C > C-A/T > T-T ≈ A-A > G-A/T > G-G[4,-11]. Second, the free energy of substrates containing mismatches can be estimated using available algorithms, such as NUPACK[12-13].

In this study, I designed 14 substrates with free energy values ranging from -39.1 kcal/mol to -28.3 kcal/mol. I then used these substrates to perform mismatch-aided DNA catalytic reactions

and studied the correlations between the free energy of the substrate and the kinetics of both catalyzed and uncatalyzed reactions.

3.2 Experimental section

3.2.1 DNA oligonucleotides

DNA oligonucleotides (oligos) were synthesized, purified, and labeled by Integrated DNA Technologies (IDT, Coralville, IA). The sequence and labeling of the used DNA oligos are listed in Tables 3.1.

3.2.2 Substrate preparation

Substrates were prepared by simple annealing without the need for purification. For the preparation of the substrate, oligo **D** was mixed with oligo **O** at a molar ratio of 1.2:1 in 10 mM TE buffer (pH 8.0) containing 14 mM $MgCl_2$ and 0.05% w/v Tween 20, with a final concentration of **O** at 5 μM. For the preparation of the reporter, oligo **L** was mixed with oligo **Q** at a molar ratio of 1:1.2 in 10 mM TE buffer (pH 8.0) containing 14 mM $MgCl_2$ and 0.05% w/v Tween 20, with a final concentration of **L** at 10 μM. The mixtures were loaded into a thermocycler for annealing using a program of 80°C for 5 minutes, cooling down to 20°C at a rate of 0.1°C every 5 seconds. The substrates were stored at 4°C prior to use.

3.2.3 Free energy of formation estimation

The free energy of the substrates, intermediates, and product was estimated through NUPACK3. The concentrations of Na^+ and Mg^{2+} were set at 50 mM and 20 mM, respectively. The free energy of all single-stranded strands is considered zero because there are not stable secondary structures in those strands.

3.2.4 Determination of rate constants

Rate constants k_1 and k_2 was determined through the following equations using elementary step 1 as an example.

$$\text{Rate} = -\frac{d[S]}{dt} = -\frac{d[C]}{dt} = \frac{d[I3]}{dt} = \frac{d[O]}{dt} = k_1[S][C] \qquad (1)$$

Where [S] is molar concentration of the substrate, [C] is molar concentration of the catalyst, and k_1 is the rate constant.

$$[S] = ([S]_0 - x); [C] = ([C]_0 - x). \qquad (2)$$

Where $[S]_0$ represents the initial concentration of the substrate, $[C]_0$ represents the initial concentration of the catalyst, and x represents amount of the substrate or catalyst that has been reacted. After substituting equation (2) into equation (1), the following equations are obtained:

$$\frac{d([S]_0 - x)}{dt} = -k_1([S]_0 - x)([C]_0 - x) \qquad (3)$$

$$-\frac{dx}{dt} = -k_1([S]_0 - x)([C]_0 - x) \qquad (4)$$

Solving of equation 4 was performed using MATLAB codes created by Zhang and Winfree[10] with minor modifications.

For accurate determination of the reaction process, 1 nM substrate, 2 nM of reporter, and 2 nM catalyst were used in a 100μL-reaction mixture. The reaction buffer consisted of 10 mM Tris-EDTA (TE) buffer (pH 8.0), 20 mM $MgCl_2$, and 0.05% w/v Tween 20. These relatively low reactant concentrations allowed the reaction to occur at a slow rate while generating a sufficient fluorescence signal for detection. The released output reacts with the reporter through TMSD to separate the fluorophore from the quencher, restoring the fluorescent signal. This reaction also prevents the reverse reaction as the output DNA is consumed by this fast reaction. The reaction progress was monitored in real-time using a fluorescence plate reader, with fluorescence detected every 3 minutes for one hour at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

Because the rate of the uncatalyzed reaction is slow, 100 nM substrate and 100 nM fuel were used in a 100μL-reaction mixture for determining k_0. Fluorescence was detected every 5 minutes for two hours. The reaction curves were then used to calculate 1/[S], and k_0 can be obtained by plotting 1/[S] against t because:

$$\frac{1}{[S]} = \frac{1}{[S]_0} + k_0 \qquad (5)$$

3.2.5 Performing the mismatch-aided DNA catalytic reaction.

Unless stated otherwise, the mismatch-aided DNA catalytic reaction was conducted using 100 μL of reaction solutions containing 25 nM substrate, 100 nM fuel, 20 nM reporter, and the desired concentration of catalyst in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. The reaction progress was monitored in real-time using a fluorescence plate reader,

with fluorescence detected every 3 minutes for one hour at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

3.2.6 Performing the two-layer mismatch-aided DNA catalytic reaction

Unless stated otherwise, the mismatch-aided DNA catalytic reaction was conducted using 100 μL of reaction solutions containing 5 nM substrate 1, 100 nM fuel 1, 20 nM substrate 2 or reporter 2 (**Q2:L2**), and 200 nM fuel 2 (**F2**) and the desired concentration of catalyst in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. The reaction progress was monitored in real-time using a fluorescence plate reader, with fluorescence detected every 3 minutes for one hour at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

3.3 Results and discussion

3.3.1 Principle of mismatch-aided non-covalent DNA amplification.

Unlike the TWJ-mediated DNA catalytic reaction which uses a TWJ to destabilize the substrate, the mismatch-aided DNA catalytic reaction introduces one or two mismatches into the substrate to achieve destabilization. The presence of a mismatch in the substrate increases its free energy, making the overall catalytic process energetically favorable. Compared to TWJ, using mismatches to modulate the free energy of the substrate is simpler and more flexible for creating substrates with varying free energy levels. However, the presence of a mismatch in the substrate leaves one nucleotide unpaired, potentially increasing the risk of leakage.

Figure 3.1 shows the scheme of the mismatch-aided DNA catalytic reaction. To simplify the explanation, I mapped each DNA strand into different functional domains labeled by numbers, where domain **n** is complementary to domain **n***. The arrow denotes the 3'-end of a DNA strand. The substate (**S**) is formed by hybridization of domains **2**, **3**, **4** of the oligo **D** and the domains **2***, **3***, **4*** of the output DNA (**O**). The oligo **D** also contains a unpaired domain **1** at its 5′-end acting as toehold, and the output DNA also has a unpaired domain **5** at its 5'-end used to react with the reporter after the output DNA is released. One or two mismatched bases are designed in the output DNA to destabilize the substrate. The fuel (**F**) shares the same domains **2***, **3***, and **4*** with the output DNA except no mismatched bases. In the absence of the catalyst (**C**), the reaction rate of the fuel with the substrate is slow because only one or two unpaired nucleotides are present in the domain **3**. In the presence of the catalyst, the catalyst can catalyze the reaction between the substrate and fuel through a multistep reaction involving two toehold exchange reactions (**Figure 3.1 B**). Specifically, the domain **1*** of the catalyst first hybridizes to the toehold **1** of the substrate, initiating the first toehold exchange reaction. Subsequent branch migration leads to the formation of intermediate ***I*** and release of the output DNA. Meanwhile, the previously blocked toehold **4** is exposed. The released output DNA then reacts with the reporter through a toehold-mediated strand displacement, restoring the fluorescence of a previously quenched fluorophore molecule. Next, the domain 4* of the fuel binds to the toehold 4, initiating the second toehold exchange reaction. This reaction then forms the product DNA and regenerates the catalyst. Consequently, the catalyst reacts with another substrate molecule, and the process repeats, leading to linear signal amplification of the catalyst.

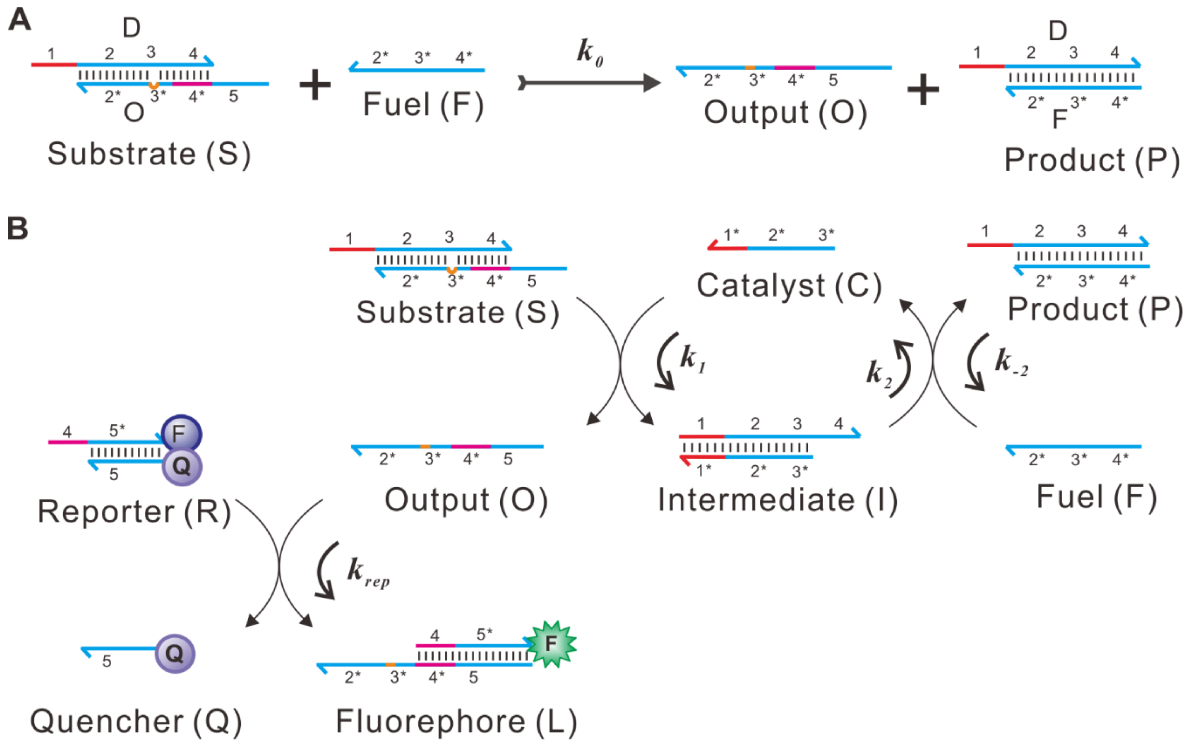


Figure 3.1 *Schemes of mismatch-aided DNA catalytic reaction.*

A) Uncatalyzed reaction. B) Pathway of the catalytic reaction.

Similar to the TWJ-mediated DNA catalytic reaction, a two-step model was used to describe the mismatch-aided DNA catalytic reaction, where two toehold exchange reactions are used as the elementary steps:

$$\text{Overall Reaction (OR): } S + F \underset{k_{-1}k_{-2}}{\overset{k_1k_2}{\rightleftarrows}} P + O$$

$$\text{Elementary reaction 1(es1): } S + C \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} I + O$$

$$\text{Elementary reaction 2(es2): } I + F \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} P + C$$

$$\text{Uncatalyzed reaction (uc): } S + F \underset{k_0}{\overset{k_0}{\rightleftarrows}} P + O$$

Signal generation: $O + R \xrightarrow{k_{rep}} E + Q$

If the mismatch is absent, DNA reactants are identical to DNA products and free energy change ΔG^0 of the overall reaction is zero. The free energy of the elementary step 1 is determined by difference in binding energy of two toeholds, **1** and **4**. Because the step 2 is a reverse reaction of the step1, the free energy of the elementary step 2 is therefore determined by difference in binding energy of two toeholds, **4** and **1**. In our design, the binding energy of the toehold **4** is lower than that of the toehold **1**. Therefore, the free energy of the elementary step 1 is positive, suggesting the step 1 is nonspontaneous, while the free energy of the elementary step 2 is negative, suggesting the step 2 is spontaneous. Therefore, the step 1 is the rate limiting step, determining the rate of the overall reaction. The introduction of a mismatch of into the substrate makes the free energy of the steps negative and does not affect the step 2. Therefore, both steps are spontaneous, thereby improving the kinetics of the catalytic reaction.

3.3.2 Design of substrates with varying free energy levels

To study correlations between the free energy of the substrate and the overall catalytic reaction, I designed substrates with different free energy levels by introducing mismatches into the output DNA. To minimize the effect of the mismatch location on the kinetics of the reactions, I introduced mismatches within the domain **3*** of the output DNA. As shown in **Table 3.1**, I mutated one or two nucleotides of the output DNA with different bases to create substrates with varying free energy levels. Six mutated output DNA strands were designed. Additionally, I designed two oligo **D** strands **D** and **D-1**, where the oligo **D-1** hybridized to the output DNA with one more base pair than the oligo **D**. The respective hybridization of **D** and **D-1** with the unmutated and the six

mutated output DNA strand results in 14 substrates. The free energy of formation of these 14 substrates estimated using NUPACK3[12] are listed in **Table 3.1**.

Table 3.1 *Sequences used for creating mismatch substrates with different free energy of formation. Nucleotide highlighted in* ***orange*** *are the mismatched bases.* ΔG^o_{fs} *was estimated using NUPACK3[12].*

	Sequence	ΔG^o_{fs} of **D:O** (kcal/mol)	ΔG^o_{fs} of **D-1:O** (kcal/mol)
D	CAT CAC CTC ACT CAA C CTA CCA C ATC TCC A 1 2 3 4		
D-1	CAT CAC CTC ACT CAA C CTA CCA C ATC TCC AC 1 2 3 4		
O-N	AGA AGG TGA GTA GTG TGG AGA T GTG GTA G GTT GAG TGA G 5 4* 3* 2*	**-36.9**	**-39.1**
O-1	AGA AGG TGA GTA GTG TGG AGA T GTG GTG G GTT GAG TGA G 5 4* 3* 2*	**-34.5**	**-36.8**
O-2	AGA AGG TGA GTA GTG TGG AGA T GTG GAA G GTT GAG TGA G 5 4* 3* 2*	**-33.5**	**-35.8**
O-3	AGA AGG TGA GTA GTG TGG AGA T GTG GTA T GTT GAG TGA G 5 4* 3* 2*	**-32.1**	**-34.4**
O-4	AGA AGG TGA GTA GTG TGG AGA T GTG ATA G GTT GAG TGA G 5 4* 3* 2*	**-31.2**	**-33.5**
O-5	AGA AGG TGA GTA GTG TGG AGA T GAG GAA G GTT GAG TGA G 5 4* 3* 2*	**-29.7**	**-32.0**
O-6	AGA AGG TGA GTA GTG TGG AGA T GAG GTA T GTT GAG TGA G 5 4* 3* 2*	**-28.3**	**-30.6**
F	TGG AGA T GTG GTA G GTT GAG TGA G 4* 3* 2*	**-36.9**	**-36.9**
C	GTG GTA G GTT GAG TGA G GTG ATG 3* 2* 1*	**-35.8**	**-35.8**
Q	5IABkFQ/ -AGA AGG TGA GTA GTG 5		
L	ATC TCC A CA CTA CTC ACC TTC T/FAM/ 4 5*		
Q2	ATC TCC A CAC TAC TCA CCT TCT CTT CCT A /IABkFQ/ 4 5* 6		
L2	/FAM/ T AGG AAG AGT AGG TGA GTA GTG T 6* 5		
F2	T AGG AAG AGA AGG TGA GTA GTG T 6* 5		

3.3.3 Correlation between free energy of formation of substrate (ΔG^o_{fS}) and rate constants of the elementary step 1 (k_1)

I then used these 14 substrates to perform the reaction of the elementary step 1 and determined the rate constant k_1 (**Table 3.2**). The substrate with the highest free energy of formation (ΔG^o_{fS}=-28.3 kcal/mol) resulted in the highest k_1 (3.32×10^6 $M^{-1}s^{-1}$), while the substate with the lowest ΔG^o_{fS} (-39.14 kcal/mol) led to the lowest k_1 (8.5×10^4 $M^{-1}s^{-1}$), which is about 39.1-fold lower than the highest k_1. I then examined the correlation between the ΔG^o_{fS} and the rate constant k_1. Interestingly, a linear relationship was observed between the ΔG^o_{fS} and rate constant k_1, suggesting that increasing the ΔG^o_{fS} proportionally increases the rate constant k_1 (**Figure 3.2 A**).

Because the same intermediate ***I*** was produced for all these step 1 reactions using these 14 substates, the free energy changes of these step 1 reactions can be calculated by subtracting the free energy of formation the substrates (ΔG^o_{fS}) from the free energy of the formation of the intermediate (ΔG^o_{fI}=-37.4 kcal/mol) (**Table 3.2**). Similarly, a linear relationship was also observed between the free energy change of the step 1 and rate constant k_1 (**Figure 3.2 B**). The step 1 reactions using these 14 substrates involve the same forward and reverse toehold domains **1** and **4**, that play a key role in determining kinetics of the forward and reverse reactions. The decrease of the free energy change of the reaction from positive to negative facilitates the forward reaction and impedes the reverse reaction, thereby increasing the rate constant. However, when the free energy change of the reaction decreases to -2 kcal/mol, the k_1/k_{-1} is over 10 and the reverse reaction becomes negligible. However, further decreasing the free energy change of the reaction continuously increases k_1, suggesting that increasing the free energy of the substrate with mismatches can also affect the rate of branch migration.

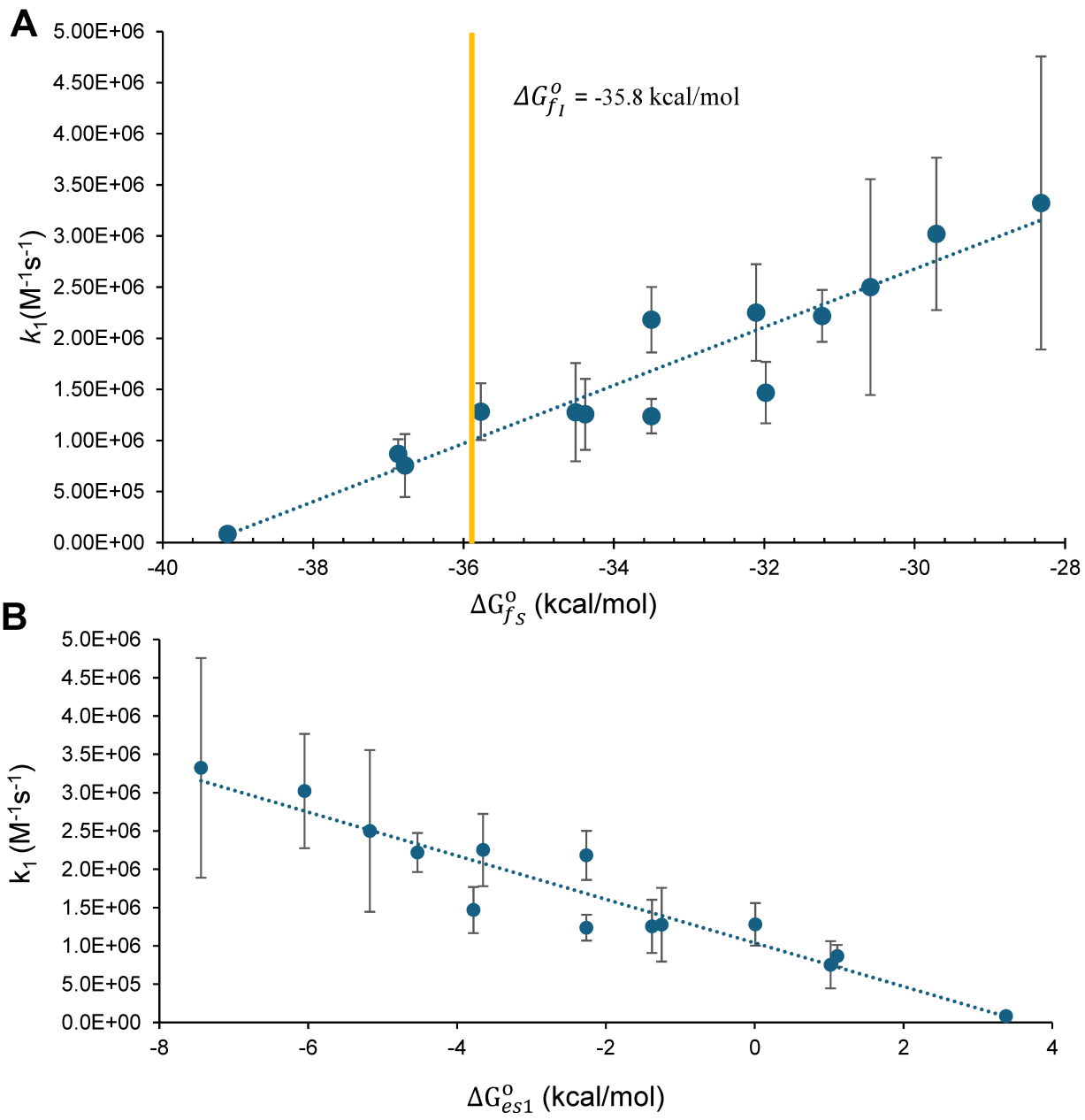


Figure 3.2 *The relationship between rate constant (k_1) of the elementary step 1, and (A) free energy of formation of the substrate($\Delta G^o_{f_S}$); (B) the free energy change of the elementary step 1 (ΔG^o_{es1}).*

The yellow line indicates the $\Delta G^o_{f_I}$ (-37.4 kcal/mol), substrates with $\Delta G^o_{f_S}$ higher than the$\Delta G^o_{f_I}$ result in a negative ΔG^o_{es1} and the elementary step one is energetically favorable.

Table 3.2 *k_1 of substrates containing different mismatches.*						
	D:O			**D-1:O**		
	ΔG^o_{fs} **(kcal/mol)**	ΔG^o_{es1} **(kcal/mol)**	k_1 **($M^{-1}s^{-1}$)**	ΔG^o_{fs} **(kcal/mol)**	ΔG^o_{es1} **(kcal/mol)**	k_1**($M^{-1}s^{-1}$)**
O-N	**-36.9**	**1.1**	**$8.7 \times 10^5 \pm 1.5 \times 10^5$**	**-39.1**	**3.4**	**$8.5 \times 10^4 \pm 4.8 \times 10^4$**
O-1	**-34.5**	**-1.2**	**$1.3 \times 10^6 \pm 4.8 \times 10^5$**	**-36.8**	**1.0**	**$7.5 \times 10^5 \pm 3.1 \times 10^5$**
O-2	**-33.5**	**-2.3**	**$2.2 \times 10^6 \pm 3.2 \times 10^5$**	**-35.8**	**0.0**	**$1.3 \times 10^6 \pm 2.8 \times 10^5$**
O-3	**-32.1**	**-3.6**	**$2.3 \times 10^6 \pm 4.7 \times 10^5$**	**-34.4**	**-1.4**	**$1.3 \times 10^6 \pm 3.5 \times 10^5$**
O-4	**-31.2**	**-4.5**	**$2.2 \times 10^6 \pm 2.5 \times 10^5$**	**-33.5**	**-2.3**	**$1.2 \times 10^6 \pm 1.7 \times 10^5$**
O-5	**-29.7**	**-6.0**	**$3.0 \times 10^6 \pm 7.5 \times 10^5$**	**-32.0**	**-3.8**	**$1.5 \times 10^6 \pm 3.0 \times 10^5$**
O-6	**-28.3**	**-7.4**	**$3.3 \times 10^6 \pm 1.4 \times 10^6$**	**-30.6**	**-5.2**	**$2.5 \times 10^6 \pm 1.1 \times 10^6$**

3.3.4 Determination of correlation between free energy of formation of substrate (ΔG^o_{fs}) and rate constants of the leakage (k_0)

Introducing mismatches into the substrate means that unpaired bases are present in the substrate, which can increase risk of leakage. To reveal the correlation between the free energy of formation of the substrate (ΔG^o_{fs}) and rate constant k_0, I used these 14 substrates to conduct uncatalyzed reactions between the substrate and fuel and determined k_0 of these reactions. Similar to k_1, the substrate with the highest free energy (-28. 3 kcal/mol) led to the highest k_0 (120 $M^{-1}s^{-1}$), while the substate with the lowest free energy (-39.1 kcal/mol) resulted in the lowest k_0 (10.6 M^{-}

$^{1}s^{-1}$) (**Table 3.3**). Therefore, increasing the $\Delta G^{o}_{f_S}$ from -39.1 kcal/mol to -28.3 kcal/mol improved k_0 by 10.6-folds. This improvement is about 3.7 times lower than that of k_1, suggesting that the increasing the $\Delta G^{o}_{f_S}$ has a smaller impact on k_0 than on k_1. A nonlinear relationship was observed between the $\Delta G^{o}_{f_S}$ and rate constant k_0, indicating as the $\Delta G^{o}_{f_S}$ increases, k_0 increases at an accelerating rate (**Figure 3.3 A**). Therefore, the increase of $\Delta G^{o}_{f_S}$ decreases the free energy change of the uncatalyzed reaction, facilitating the occurrence of the reaction and increasing the risk of the leakage (**Figure 3.3 B**).

I then calculated the ratio of k_1/k_0 for each substrate and examined the correlation between the free energy of formation of the substrate ($\Delta G^{o}_{f_S}$) and the k_1/k_0 ratio (**Table 3.4 and Figure 3.4**). Because k_1 and k_0 are associated with signal amplification and background of the catalytic reaction, the k_1/k_0 ratio relates to the signal-to-background ratio. The k_1/k_0 ratio for the substrate with the lowest $\Delta G^{o}_{f_S}$ (-39.1 kcal/mol) is significantly lower than that of the substrate with $\Delta G^{o}_{f_S}$ equal to that of the product DNA. Increasing the $\Delta G^{o}_{f_S}$ above that of the product ($\Delta G^{o}_{f_P}$) significantly improved the k_1/k_0 ratio (**Figure 3.4**). Except for two substrates having a $\Delta G^{o}_{f_S}$ of -35.8 kcal/mol and -34.5 kcal/mol, other substrates with free energy levels lower than -28.3 kcal/mol showed k_1/k_0 ratios without significant differences. However, further increasing the $\Delta G^{o}_{f_S}$ to -28.3 kcal/mol significantly decreased the k_1/k_0 ratio. These results suggest that increasing the $\Delta G^{o}_{f_S}$ by about 7 kcal/mol higher than that of the product DNA does not significantly change the k_1/k_0 ratio.

Table 3.3 *k_0 of substrate containing different mismatches.*						
	D:O			**D-1:O**		
	ΔG^o_{fs} **(kcal/mol)**	ΔG^o_{uc} **(kcal/mol)**	k_0 **($M^{-1}s^{-1}$)**	ΔG^o_{fs} **(kcal/mol)**	ΔG^o_{uc} **(kcal/mol)**	k_0 **($M^{-1}s^{-1}$)**
O-N	**-36.9**	**0.0**	**31.0 ± 1.1**	**-39.1**	**2.3**	**10.6 ± 0.8**
O-1	**-34.5**	**-2.4**	**43.7 ± 0.5**	**-36.8**	**-0.1**	**15.4 ± 0.4**
O-2	**-33.5**	**-3.4**	**43.7 ± 2.5**	**-35.8**	**-1.1**	**17.0 ± 0.8**
O-3	**-32.1**	**-4.8**	**54.8 ± 3.3**	**-34.4**	**-2.5**	**26.6 ± 2.1**
O-4	**-31.2**	**-5.6**	**54.7 ± 7.1**	**-33.5**	**-3.4**	**19.5 ± 0.6**
O-5	**-29.7**	**-7.2**	**77.5 ± 5.4**	**-32.0**	**-4.9**	**26.7 ± 3.3**
O-6	**-28.3**	**-8.6**	**120.0 ± 29.0**	**-30.6**	**-6.3**	**50.7 ± 12.6**

Table 3.4 *k_1/k_0 of substrate containing different mismatches.*						
	D:O			**D-1:O**		
	ΔG^o_{fs} **(kcal/mol)**	ΔG^o_{es1} **(kcal/mol)**	k_1/k_0	ΔG^o_{fs} **(kcal/mol)**	ΔG^o_{es1} **(kcal/mol)**	k_1/k_0
O-N	**-36.9**	**1.1**	**$2.8 \times 10^4 \pm 2.9 \times 10^3$**	**-39.1**	**3.4**	**$8.0 \times 10^3 \pm 1.2 \times 10^3$**
O-1	**-34.5**	**-1.2**	**$2.9 \times 10^4 \pm 1.2 \times 10^4$**	**-36.8**	**1.0**	**$4.9 \times 10^4 \pm 3.3 \times 10^4$**
O-2	**-33.5**	**-2.3**	**$5.0 \times 10^4 \pm 5.3 \times 10^3$**	**-35.8**	**0.0**	**$7.5 \times 10^4 \pm 7.2 \times 10^3$**
O-3	**-32.1**	**-3.6**	**$4.1 \times 10^4 \pm 4.9 \times 10^3$**	**-34.4**	**-1.4**	**$4.7 \times 10^4 \pm 7.6 \times 10^3$**
O-4	**-31.2**	**-4.5**	**$4.1 \times 10^4 \pm 1.1 \times 10^4$**	**-33.5**	**-2.3**	**$6.3 \times 10^4 \pm 3.6 \times 10^3$**
O-5	**-29.7**	**-6.0**	**$3.9 \times 10^4 \pm 4.9 \times 10^3$**	**-32.0**	**-3.8**	**$5.5 \times 10^4 \pm 1.4 \times 10^4$**
O-6	**-28.3**	**-7.4**	**$2.8 \times 10^4 \pm 1.5 \times 10^4$**	**-30.6**	**-5.2**	**$4.9 \times 10^4 \pm 2.6 \times 10^4$**

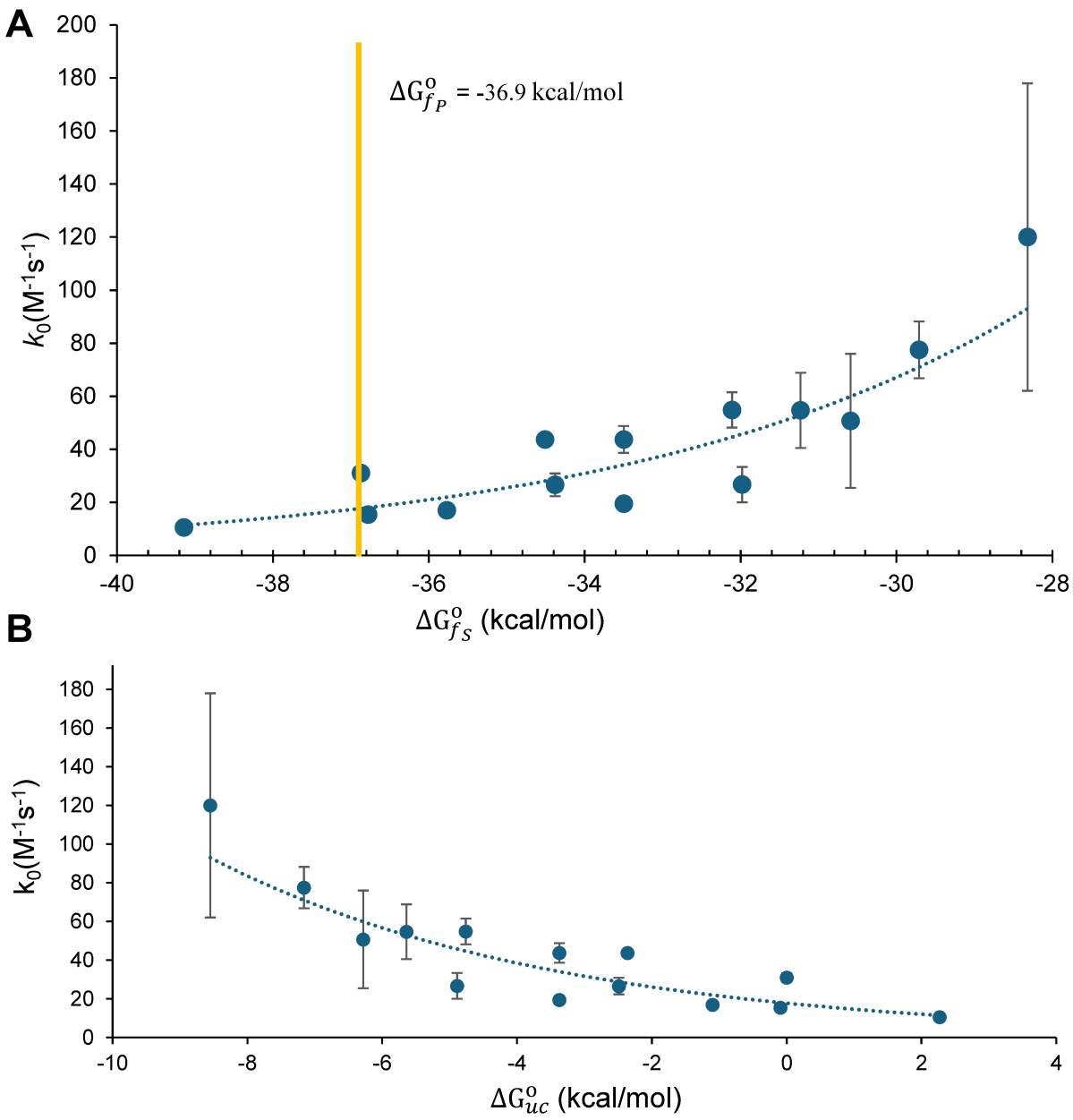


Figure 3.3 *The relationship between rate constant of the unanalyzed reaction (k_0) and A) the free energy of formation of the substrates ($\Delta G^o_{f_S}$); B) the free energy change of the unanalyzed reaction (ΔG^o_{uc}).*

The yellow line indicates the $\Delta G^o_{f_P}$ (-36.9 kcal/mol), substrates with $\Delta G^o_{f_S}$ higher than the$\Delta G^o_{f_P}$ result in a negative ΔG^o_{UC} and the uncatalyzed reaction is energetically favorable.

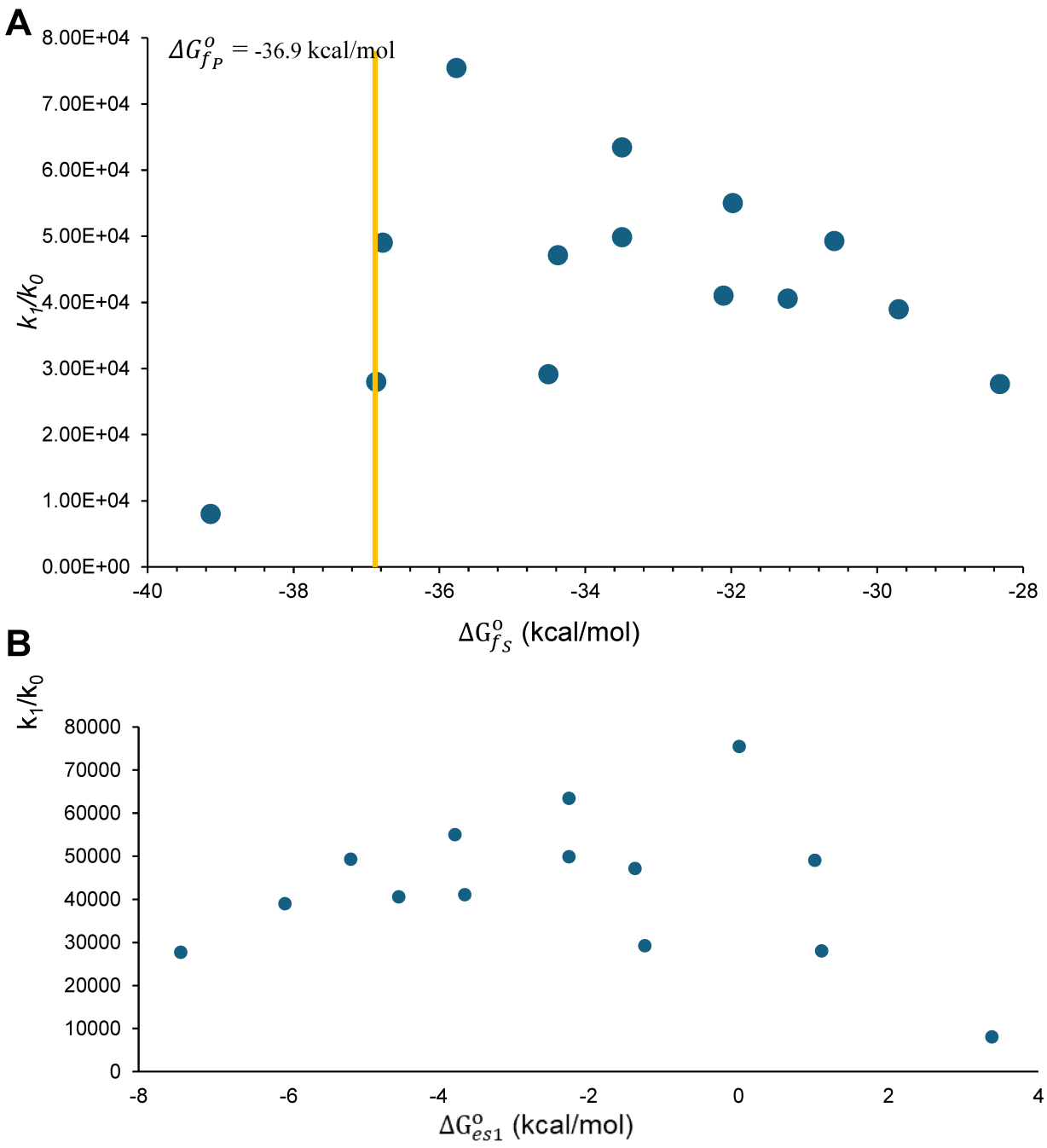


Figure 3.4 *Relationship between k_1/k_0 and A) the free energy of formation of substrates($\Delta G^o_{f_S}$); B) the free energy change of the elementary step 1 (ΔG^o_{es1}).*

The yellow line indicates the $\Delta G^o_{f_P}$ (-36.9 kcal/mol), substrates with $\Delta G^o_{f_S}$ higher than the$\Delta G^o_{f_P}$ result in a negative ΔG^o_{OR}.

3.3.5 Comparison of the mismatch-containing substrate with the non-mismatch substrate for DNA catalysis.

I then selected the substrate formed by hybridization of **D-1** and **O-3** to perform the mismatch-aided DNA catalytic reaction. The mismatch-aided DNA catalytic reaction was compared with a DNA catalytic reaction using a substrate without mismatch (also called non-mismatch substrate) for signal amplification. Two reactions are prepared containing 25 nM of mismatch substrate or non-mismatch substrate, 100 nM fuel, 20 nM reporter, and 100 pM catalyst in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. The reactions were monitored with real-time fluorescence detection for one hour. The fluorescence signal was then converted to the amount of the output DNA using a calibration curve established with a series of standard solutions containing varying concentrations of the output DNA and 20 nM reporter. The mismatch-aided DNA catalytic reaction resulted in approximately 4.7 nM output DNA from 100 pM catalyst, representing 47-fold signal amplification in a one-hour reaction. In contrast, the DNA catalytic reaction using the non-mismatch substrate produced about 0.6 nM product, equivalent to 6-fold signal amplification. Therefore, compared to the DNA catalytic reaction using the non-mismatch substrate, the mismatch-aided DNA catalytic reaction improved the amplification efficiency by 7.8 folds.

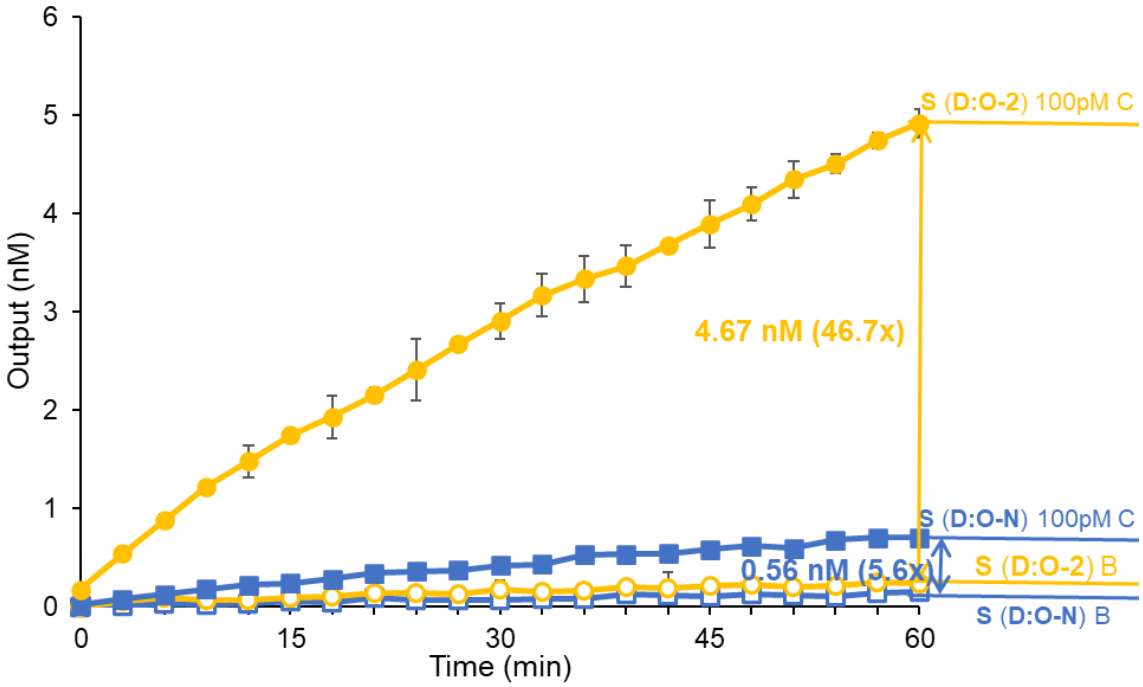


Figure 3.5 *Comparison of the mismatch substrate with the non-mismatch substrate for amplification efficiency.* ***C*** *denotes the catalyst and* ***B*** *denotes the blank.*

3.3.6 Using mismatch-aided DNA catalytic reaction as an assay for amplified detection of the catalyst.

Like the TWJ-mediated DNA catalytic reaction, the mismatch-aided DNA catalytic reaction can be directly used as an assay for the amplified detection of the catalyst. To achieve high detection sensitivity, three key parameters affecting signal amplification and background of the catalytic reaction were studied, including Mg^{2+} concentration, substrate concentration, and fuel concentration. 100 pM of the catalyst was used for these studies. The Mg^{2+} concentration can affect DNA hybridization including the binding of the toehold binding domain to the toehold, thereby kinetics of the catalytic reactions. Increasing the Mg^{2+} concentration from 10 mM to 40 mM resulted in a continuous increase in signal amplification, and it also slightly increased background

(**Figure 3.6 A**). The best signal-to-background ratio was observed when Mg^{2+} was at 20 mM (**Figure 3.6 B**).

I also increased the substrate concentration from 5 nM to 20 nM and found that increasing the substrate concentration led to a significant increase in signal amplification, but a smaller increase in background (**Figure 3.7 A**). 20 nM substrate showed the best signal-to-background ratio (**Figure 3.7 B**). By contrast, increasing the fuel concentration from 50 nM to 200 nM resulted in a relatively smaller effect on the signal amplification, which suggested that the elementary step1 is a rate limiting step under these conditions (**Figure 3.8 A**). The 100 nM fuel led to the best signal-to-background ratio (**Figure 3.8 B**).

I then used the optimized conditions to determine the dynamic range and limit of detection (LOD) of the catalytic reaction as an assay. The optimized conditions are 20 nM substrate, 100 nM fuel, 20 nM reporter in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. A linear dynamic range from 5 pM to 500 pM of catalyst was obtained. The assay was able to detect as little as 2 pM of the catalyst within 1 hour at room temperature without the need for any enzymes. Therefore, the mismatch-aided DNA catalytic reaction shows the similar dynamic range and LOD to the TWJ-mediated DNA catalytic reaction.

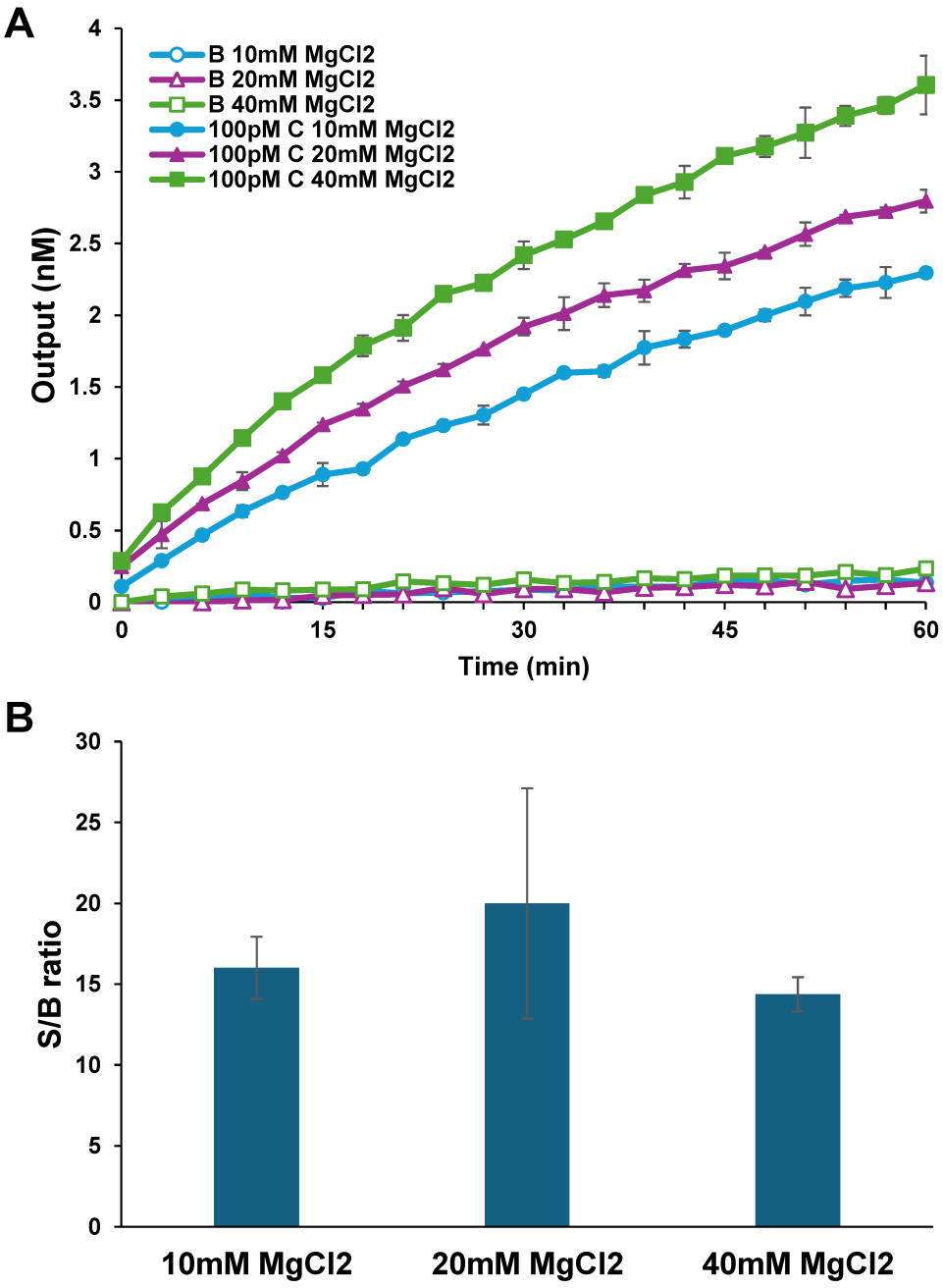


Figure 3.6 *Impact of the $MgCl_2$ concentration on the signal amplification.*

(A) Amplification curves, where the blank or background samples are denoted as ***B*** *and* ***C*** *denotes the catalyst; (B) Signal to background ratios.*

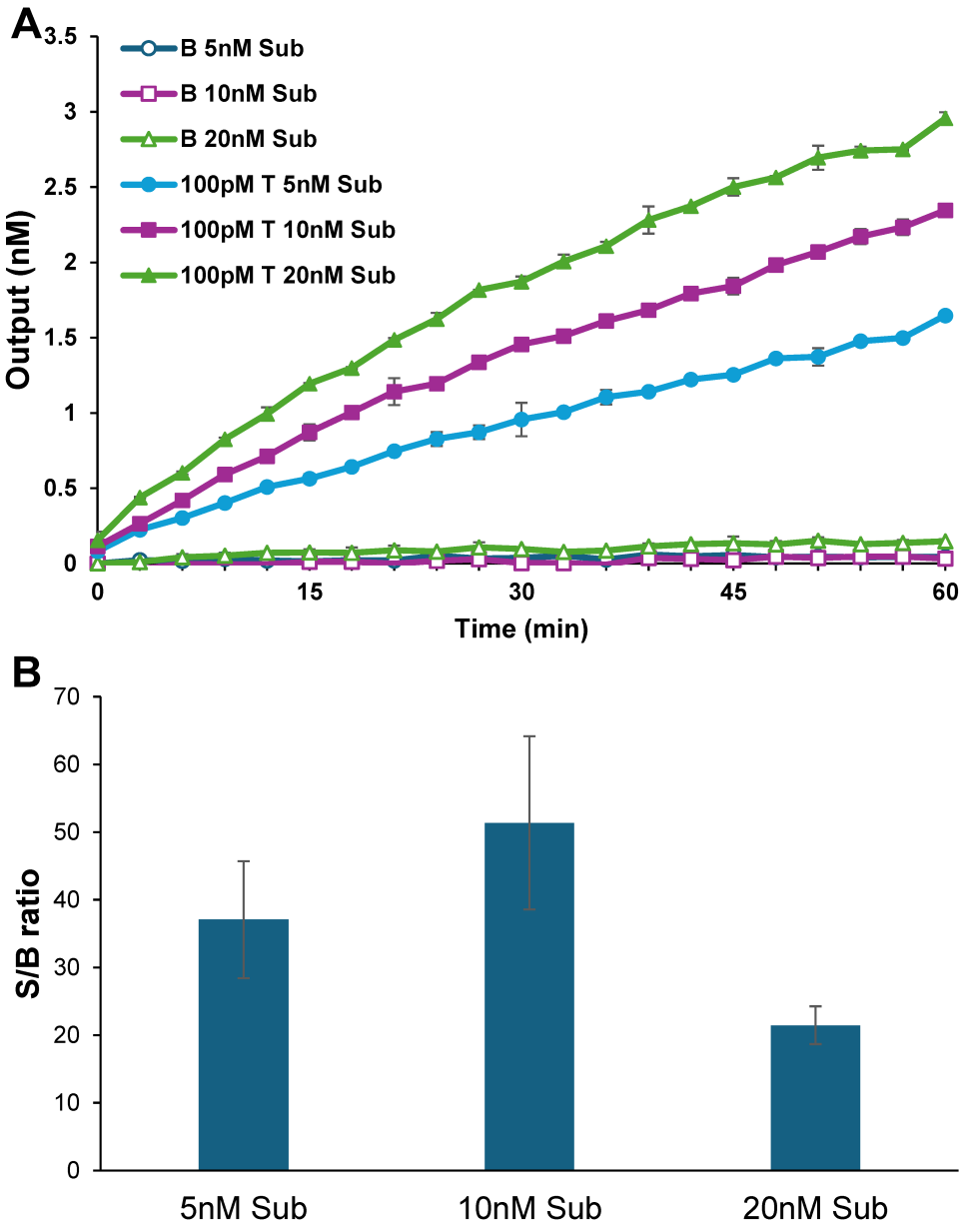


Figure 3.7 *Impact of the concentration of substrate on the signal amplification.*

(A) Amplification curves, where the blank samples are denoted as ***B****, and the catalyst is denoted as target* ***T****; (B) Signal to background ratios.*

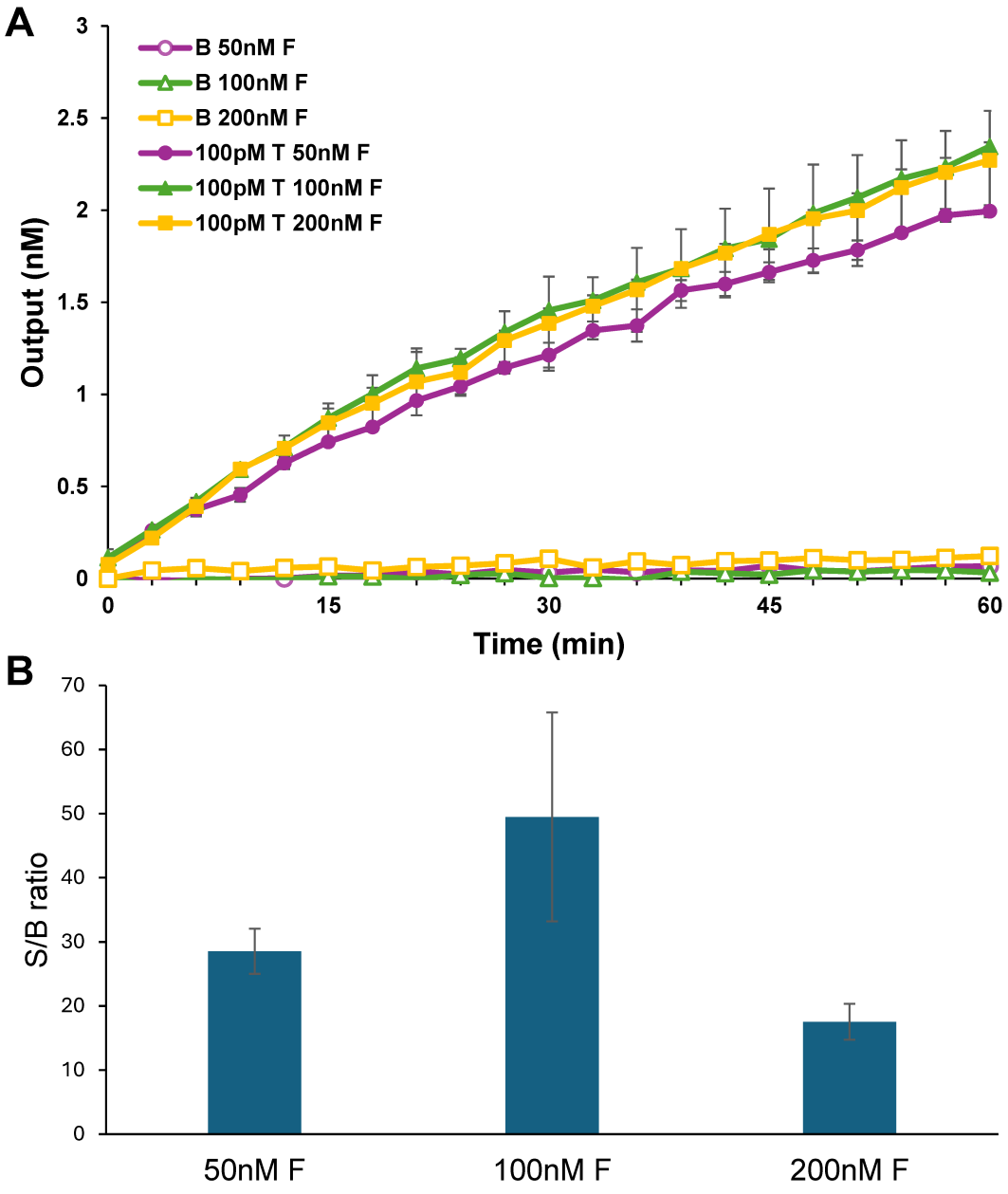


Figure 3.8 *Optimization of [F] for mismatch-aided DNA amplification assay for detection of C. (A) Amplification curves, where the blank samples are denoted as* ***B****, and the catalyst is denoted as target* ***T.*** *(B) Signal to background ratio.*

.

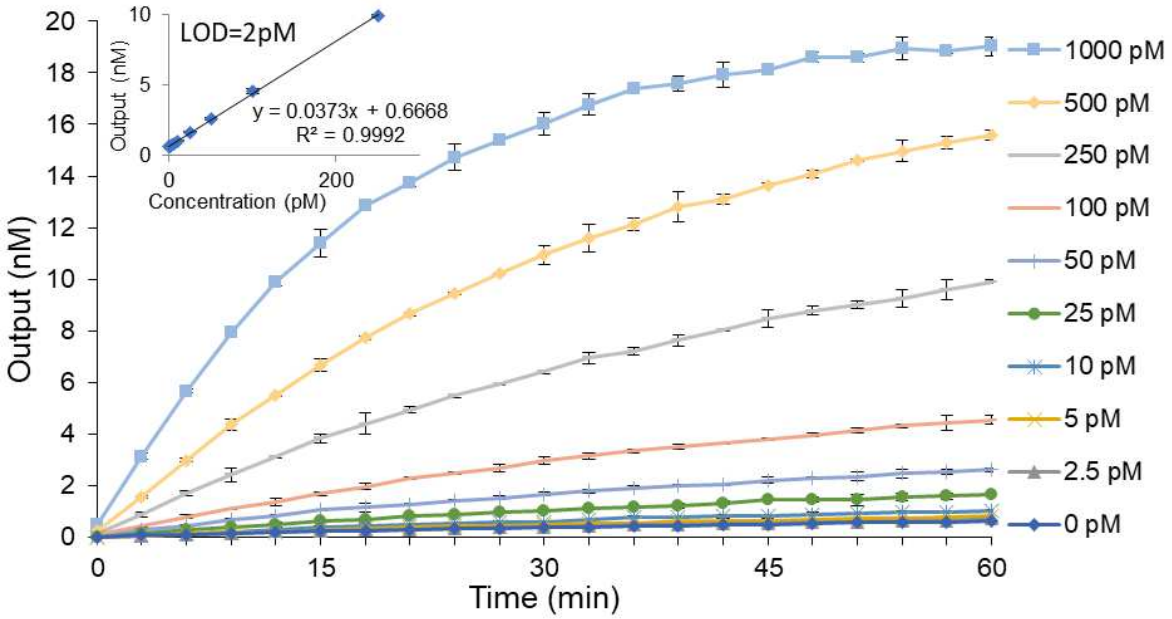


Figure 3.9 *Dynamic range and limit of detection of the mismatch-aided DNA catalytic reaction.*

3.3.7 Construction of a two-layer mismatch-aided DNA catalytic reaction

To further increase signal amplification, I also developed a two-layer mismatch-aided DNA catalytic reaction system, which uses a catalyst to initiate the first mismatch-aided DNA catalytic reaction, generating a layer 1 output DNA that then serves as the catalyst for the second mismatch-aided DNA catalytic reaction (**Figure 3.10**). As a result, one catalyst molecule can result in multiple layer 1 output DNA strands through the layer 1 DNA catalytic reaction, and each layer 1 output DNA strand further generates multiple layer 2 output DNA strands through the layer 2 DNA catalytic reaction, achieving tandem amplification. Specifically, the catalyst catalyzes the reaction of the layer 1 substrate with the layer 1 fuel, generating the layer 1 output DNA. Meanwhile, the catalyst is regenerated for catalyzing the next cycle of layer 1 reaction and each cycle of the layer 1 reaction can generate a layer 1 output DNA strand. This output DNA then acts as the catalyst to catalyze the reaction between the layer 2 substrate and layer 2 fuel, generating layer 2 output DNA.

The layer 2 output DNA further reacts with the reporter, lighting up the fluorescence of a fluorophore molecule. Similarly, the layer 2 DNA catalytic reaction is also recycled. Finally, thousands of fluorophore molecules can be turned on in response to a single catalyst molecule.

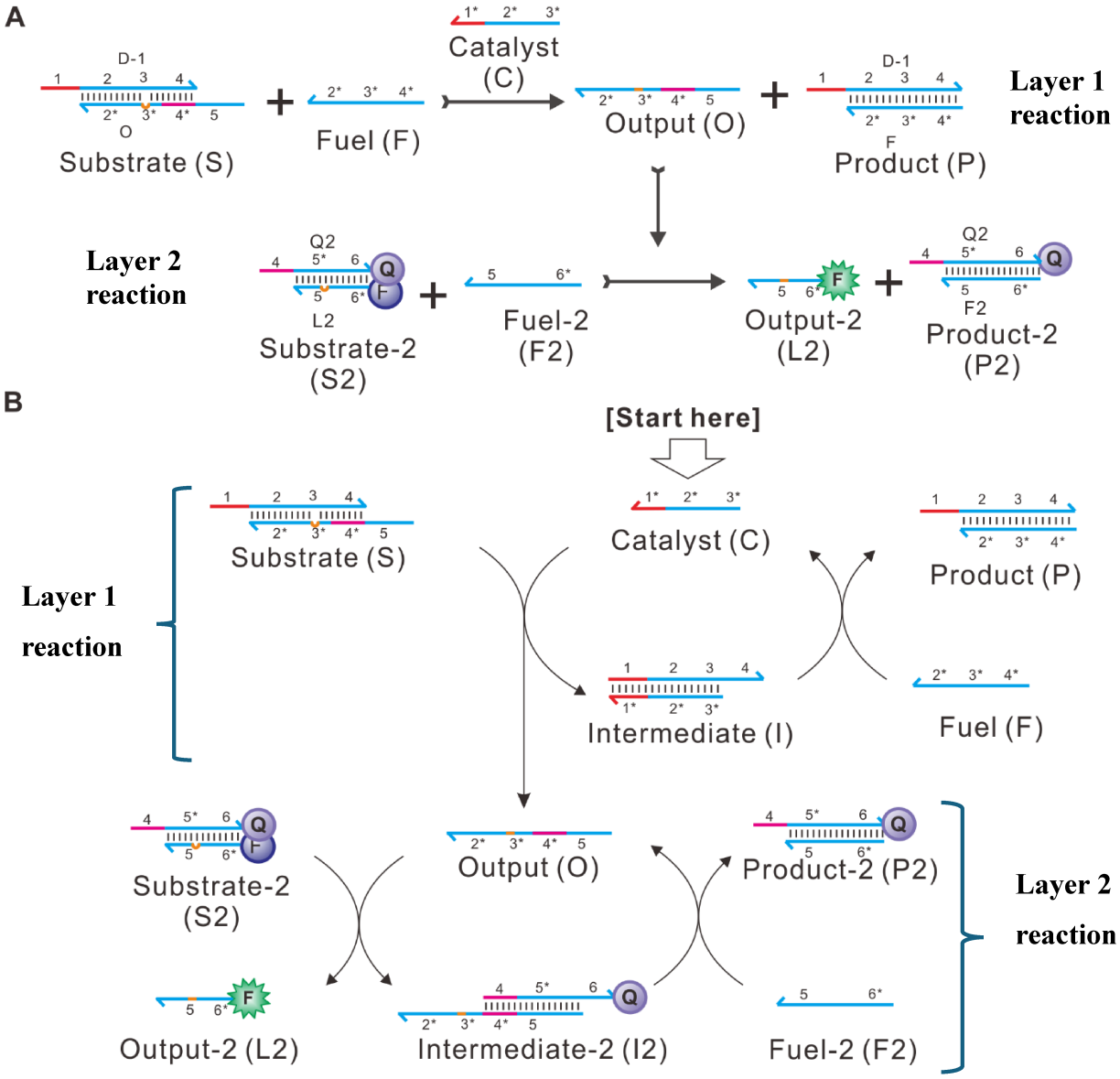


Figure 3.10 *Schemes of two-layer mismatch-aided DNA catalytic reaction.*

(A) The overall reaction; (B) the reaction pathway of the two-layered mismatch-aided DNA catalytic reaction.

I first examined the amplification efficiency of the two-layer catalytic reaction, using 10 pM catalyst and conducting the reaction for three hours (**Figure 3.11**). A 370-fold signal amplification was observed after one hour, 1170-fold after two hours, and 2060-fold after three hours. The amplification efficiency of the two-layer mismatch-aided DNA catalytic reaction is slightly better than that of the two-layer TWJ-mediated DNA catalytic reaction. However, a significantly higher background was observed for the two-layer mismatch-aided DNA catalytic reaction, which is probably due to the fact that this two-layer reaction does not have an additional step to translate the output DNA into the catalyst for the layer 2 reaction.

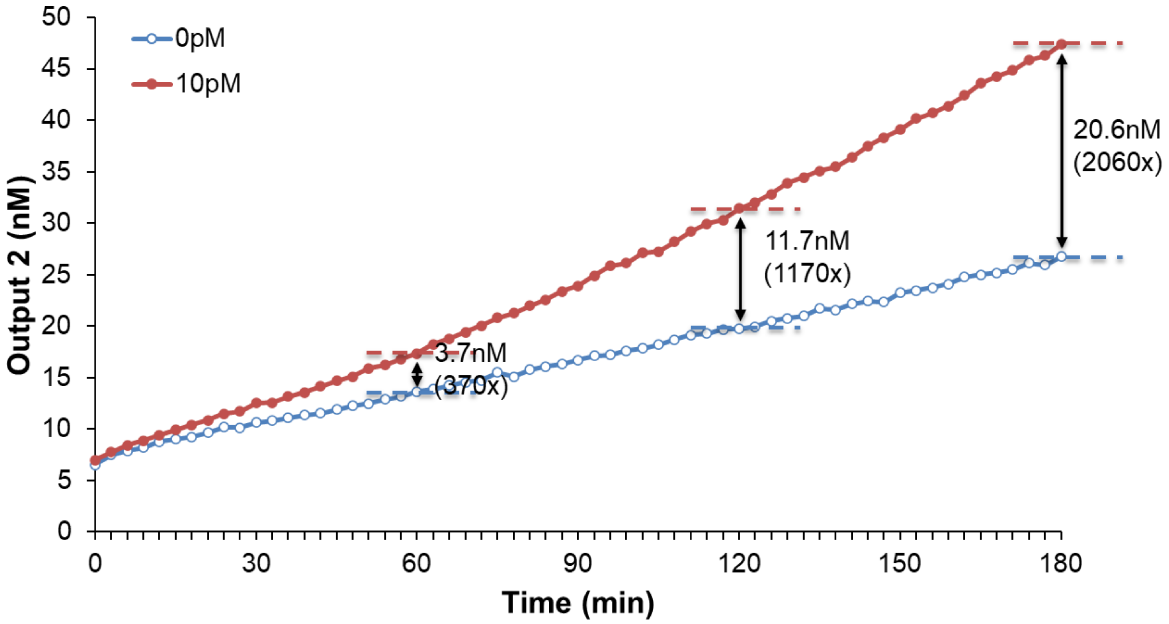


Figure 3.11 *Amplification efficiency of the two-layer mismatch-aided DNA catalytic reaction.*

I then used the two-layer catalytic reaction to analyze solutions containing varying concentrations of the catalyst (**Figure 3.12**). A dynamic range between 1 pM and 100 pM was obtained, and the reaction was able to detect as low as 0.5 pM catalyst.

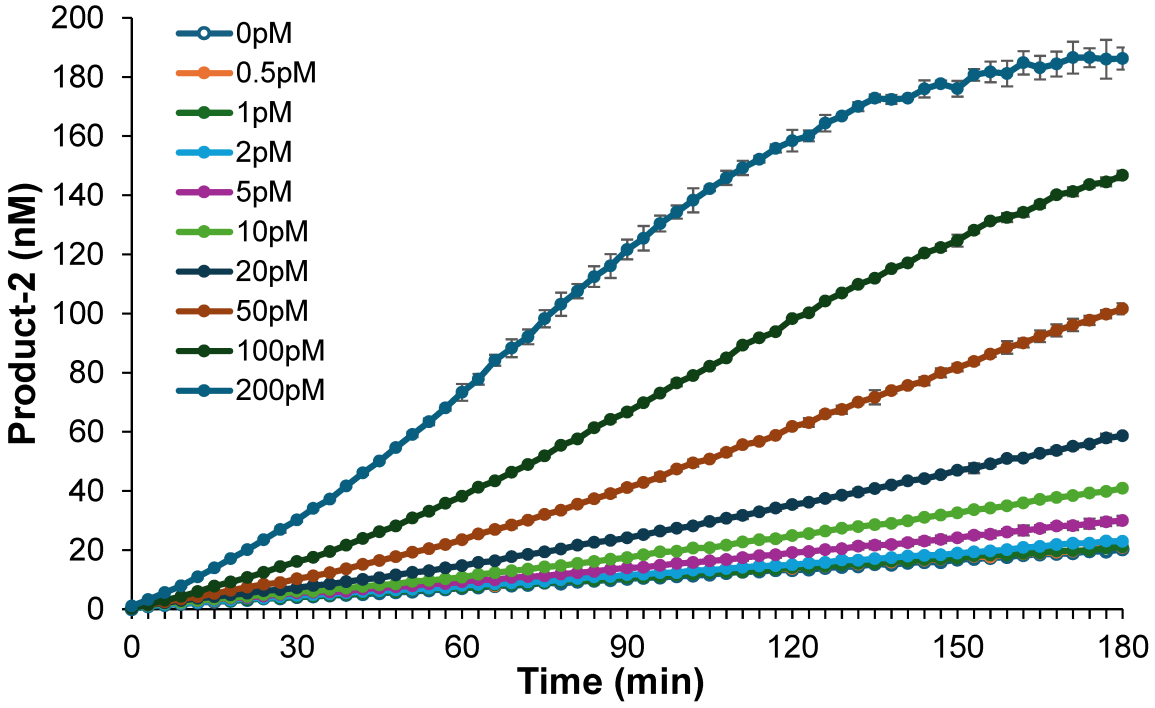


Figure 3.12 *Amplification curves of the two-layer mismatch-aided DNA catalytic reaction for detection of varying concentration of the catalyst.*

3.4 Conclusion

I developed a simple and flexible mismatch-aided DNA catalytic reaction, which is particularly suitable for studying correlations between free energy of the substrate and rate constants of catalyzed and uncatalyzed reactions. By strategically manipulating the location, base, and number of mismatches, I achieved effective modulation of the free energy of the substrate and, consequently, free energy changes of elementary steps and overall reaction. 14 substrates were designed with free energies ranging from -39.1 kcal/mol to -28.3 kcal/mol. These substrates were then used to conduct the mismatch-aided DNA catalytic reactions, and rate constant k_1 of the elementary step 1 and rate constant k_0 of the uncatalyzed reactions were determined for these

substrates to examine the relationship between between free energy of the substrate and rate constants k_1 *and* k_0.

Interestingly, a linear relationship was observed between the free energy of the substate and rate constant k_1, suggesting that increasing the free energy of the substrate proportionally increases the rate constant k_1. In contrast, A nonlinear relationship was observed between the free energy of the substate and rate constant k_0, indicating as the free energy of the substrate increases, k_0 increases at an accelerating rate. Additionally, increasing the free energy of the substrate from -39.1 kcal/mol to -28.3 kcal/mol has a less impact on k_0 than on k_1. Increasing the free energy of the substrate by about 7 kcal/mol higher than that of the product DNA does not significantly change the k_1/k_0 ratio.

I further tested the amplification efficiency of the mismatch-aided DNA catalytic reaction using a mismatch substrate with a free energy of -34.4 kcal/mol. The mismatch-aided DNA catalytic reaction exhibits an amplification efficiency 7.4 times higher than the toehold exchange catalytic reaction. When used as an assay for detection of the catalyst, the catalytic reaction can detect as low as 2 pM catalyst at room temperature without the need for any enzymes. Additionally, a two-layer mismatch-aided catalytic reaction was developed to further improve amplification efficiency. This two-layer reaction exhibited a 2060-fold signal amplification in a three-hour reaction, lowering the limit of detection to 0.5 pM for the catalyst.

The findings of this chapter reveal the relationship between the free energy change and kinetics of the toehold exchange catalytic reactions, providing valuable suggestions for designing new toehold exchange catalytic reactions.

3.5 References

1. Leontis, N. B., Kwok, W., Newman, J. S. (1991). Stability and structure of three-way DNA junctions containing unpaired nucleotides. Nucleic Acids Res. 19, 759.
2. Carr, C. E., Marky, L. A. (2017). Investigation of the melting behavior of DNA three-way junctions in the closed and open states. Biophys J. 113, 529.
3. SantaLucia Jr. J, Hicks D. (2004). The thermodynamics of DNA structural motifs. Annu Rev Biophys. 33, 415.
4. Rossetti, G., Dans, P. D., Gomez-Pinto, I., Ivani, I., Gonzalez, C., & Orozco, M. (2015). The structural impact of DNA mismatches. Nucleic Acids Res. 43(8), 4309.
5. Kundert, K., & Fraser, J. S. (2020). DNA-binding proteins meet their mismatch. Nature. 587, 494.
6. Das, G., & Lyngdoh, R. H. D. (2012). Role of wobble base pair geometry for codon degeneracy: purine-type bases at the anticodon wobble position. J Mol Model. 18(8), 3805.
7. Mathews, C. K., & Van Holde, K. E., & Appling, D. R., & Anthony-Cahill, S. J. (2012). Biochemistry (4th ed.). Toronto: Prentice Hall. p. 1181.
8. Allawi, H. T., & SantaLucia Jr., J., (1998). Nearest-neighbor thermodynamics of internal A•C mismatches in DNA: Sequence dependence and pH effects. Biochem. 37(26), 9435.
9. Peyret, N., Seneviratne, P. A., Allawi, H. T., & SantaLucia, J., Jr. (1999). Nearest-neighbor thermodynamics of deoxyribo-oligonucleotides with single base bulges. Nucleic Acids Res. 27(19), 3718.
10. Allawi, H. T., & SantaLucia, J., Jr. (1997). Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A•A, C•C, G•G, and T•T mismatches. Biochem. 36(34), 10581.
11. Yakovchuk, P., Protozanova, E., & Frank-Kamenetskii, M. D. (2006). Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. Nucleic Acids Res. 34(2), 564.
12. Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., & Pierce, N. A. (2011). NUPACK: analysis and design of nucleic acid systems. J Comput Chem. 32(1), 170.

13. Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M., & Pierce, N. A. (2017). Constrained multistate sequence design for nucleic acid reaction pathway engineering. J Am Chem Soc. 139(4), 1788.

4. Chapter 4: Modified nucleotide-assisted DNA catalytic reaction

4.1 Introduction

Studies described in chapters 2 and 3 revealed that increasing the free energy of formation of the substrate (ΔG^o_{fs}) makes the free energy change of the overall catalytic reaction (ΔG^o_{OR}) negative, thereby enhancing the kinetics and amplification efficiency of the catalytic reaction. Additionally, the ΔG^o_{fs} is associated with the kinetics of catalyzed and uncatalyzed (leakage) reactions. Specifically, increasing the ΔG^o_{fs} proportionally increases the rate constant k_1 of the elementary step 1, while the rate constant k_0 of the leakage increases at an accelerating rate as the free energy of the substrate increases. Furthermore, I also observed that the amplification efficiency of the TWJ-mediated DNA catalytic reaction was lower than that of mismatch-assisted DNA catalytic reaction when the free energy of the TWJ substrate was higher than that of the mismatch substrate. These results suggested the structure of the substrate might also impact the kinetics and amplification efficiency of the catalytic reaction, which requires more studies to elucidate this effect.

Advances in chemical synthesis of oligonucleotides now allow us to modify nucleotides by chemically altering the bases, sugars, and phosphates of DNA[1, 2]. Modification of the oligonucleotide is achieved by directly incorporating the modified nucleotide at the desired position within the oligonucleotide[3-5]. These modifications can enhance or weaken the thermostability of the DNA duplex. Modifications that destabilize the DNA duplex can be used to increase the ΔG^o_{fs}. For example, the following three modifications can reduce the stability of the

duplex: the abasic (AB) site, where the purine or pyrimidine base is removed[1, 2, 6-8]; the anucleotide (AN) site, where instead of the pentose sugar and base, there are only the C3, C4, and C5 carbons[3-5]; and the phosphorothioate (PS) linkage, where one of the oxygens connected to the phosphate is replaced with sulfur. Compared to TWJ structure and mismatches, these modifications do not generate junctions or bulges in the DNA duplex. In this study, I incorporated these three modifications into DNA and generated five types of substates. I then determined the free energy of these substrates and revealed how these modifications impact the free energy of the substrate[9-11]. I further used these substrates to construct catalytic reactions, termed modified nucleotide-assisted DNA catalytic reactions and studied the kinetics of the catalytic reactions.

4.2 Experimental section

4.2.1 DNA Oligonucleotide

DNA oligonucleotides (oligos) were synthesized, purified, and labeled by Integrated DNA Technologies (IDT, Coralville, IA). The sequence and labeling of the used DNA oligos are listed in Tables 4.1, Table 4.2 and Table 4.3.

4.2.2 Substrate preparation

Substrates were prepared by simple annealing without the need for purification. For the preparation of the substrate, oligo **D** was mixed with oligo **O** at a molar ratio of 1.2:1 in 10 mM TE buffer (pH 8.0) containing 14 mM $MgCl_2$ and 0.05% w/v Tween 20, with a final concentration of **O** at 5 μM. For the preparation of the reporter, oligo **L** was mixed with oligo **Q** at a molar ratio of 1:1.2 in 10 mM TE buffer (pH 8.0) containing 14 mM $MgCl_2$ and 0.05% w/v Tween 20, with a

final concentration of **L** at 10 μM. For the preparation of the translator, oligo **P** was mixed with oligo **C** at a molar ratio of 1.2:1 in 10 mM TE buffer (pH 8.0) containing 14 mM $MgCl_2$ and 0.05% w/v Tween 20, with a final concentration of **C** at 5 μM. The mixtures were loaded into a thermocycler for annealing using a program of 80°C for 5 minutes, cooling down to 20°C at a rate of 0.1°C every 5 seconds. For the translator, following the annealing process, the annealed translator is mixed with a blocker oligo nucleotide in 10 mM TE buffer (pH 8.0) containing 14 mM $MgCl_2$ and 0.05% w/v Tween 20, with a final concentration of **C** at 1 μM. The substrates were stored at 4°C prior to use.

4.2.3 Experimental determination of free energy of formation

The free energy of formation of substrates, intermediates, and products was calculated using the following equations:

$$\Delta G = \Delta H - T\Delta S \qquad (1)$$

$$\Delta G = -RT\ln K \qquad (2)$$

Where ΔH and ΔS are the enthalpy and entropy, respectively. T is the temperature in Kelvin (K). R value is the ideal gas constant. K is equilibrium constant of the system. Combining the two formulas into the following equation:

$$\Delta H - T\Delta S = -RT\ln K \qquad (3)$$

Moving T to one-side:

$$1/T = \frac{\Delta S - R\ln K}{\Delta H} \qquad (4)$$

When dsDNA is denatured with heating, it dissociates into ssDNA. The melting temperature (T_m) is the temperature at which 50% of the dsDNA is dissociated into ssDNA. Using our substrate **S** (**D**:**O**) that denatures into **O** and **D** as an example:

$$O + D \leftrightarrow O{:}D \qquad (5)$$

At the T_m, the reaction is at equilibrium and the equilibrium constant *K* is as following:

$$K = \frac{[O{:}D]}{[O][D]} = \frac{2\alpha}{(1-\alpha)^2 Ct} \quad (6)$$

Where the *Ct* is the total strand concentration within the system, and α is the molar fraction of DNA that is double-stranded. For example, if dsDNA is at 2 mM, *Ct* is 4mM. Therefore, at T_m, α=1/2.

Substitute equation 6 into equation 3:

$$\frac{1}{Tm} = \frac{\Delta S - R\ln(2\alpha\,/\,((1-\alpha)^2 Ct))}{\Delta H} = \frac{\Delta S + R\ln(Ct/4)}{\Delta H} \qquad (7)$$

Rearrange the equation:

$$\frac{1}{Tm} = \ln(Ct/4)\frac{R}{\Delta H} + \frac{\Delta S}{\Delta H} \qquad (8)$$

After determination of the T_m of the dsDNA at different concentrations, $1/T_m$ is plotted against ln(*Ct*/4). The slope is R/ΔH, and the Y-intercept is ΔS/ΔH. Therefore, ΔH and ΔS can be then calculated. ΔG_f^o at specific temperatures such as 25°C is determined by substituting ΔH and ΔS into equation 1.

SYBR Green, a double-stranded DNA binding dye, was used to determine the melting temperature of dsDNA. SYBR Green is fluorescent when binding to the minor groove of dsDNA and loses its fluorescence when dsDNA denatures into ssDNA. The reaction solutions were prepared containing 0.2 μM to 5 μM dsDNA duplex in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$, 1× SYBR Green dye, and 0.1× ROX reference dye. The reaction solutions were loaded

into the QuantStudio™ 3 Real-Time PCR System for melting curve analysis. The melting curve analysis started at 50°C and slowly increased the temperature at a rate of 0.01°C per second until it reached 80°C. Fluorescence was measured every 5 seconds, with an excitation wavelength of 470 nm and an emission wavelength of 520 nm. The melting temperature was determined from the first derivative of the melting curve.

4.2.4 Determination of rate constant

Rate constants k_1 and k_2 was determined through the following equations using elementary step 1 as an example.

$$\text{Rate} = -\frac{d[S]}{dt} = -\frac{d[C]}{dt} = \frac{d[I3]}{dt} = \frac{d[O]}{dt} = k_1[\text{S}][\text{C}] \qquad (9)$$

Where [S] is molar concentration of the substrate, [C] is molar concentration of the catalyst, and k_1 is the rate constant.

$$[S] = ([S]_0 - x); [C] = ([C]_0 - x) \qquad (10)$$

Where $[\text{S}]_0$ represents the initial concentration of the substrate, $[\text{C}]_0$ represents the initial concentration of the catalyst, and x represents amount of the substrate or catalyst that has been reacted. After substituting equation 10 into equation 9, the following equations are obtained:

$$\frac{d([S]_0 - x)}{dt} = -k_1([S]_0 - x)([C]_0 - x) \qquad (11)$$

$$-\frac{dx}{dt} = -k_1([S]_0 - x)([C]_0 - x) \qquad (12)$$

Solving of equation 12 was performed using MATLAB codes created by Zhang and Winfree[10] with minor modifications.

For accurate determination of the reaction process, 1 nM substrate, 2 nM reporter and 2 nM catalyst were used in a 100µL-reaction mixture. The reaction buffer consisted of 10 mM Tris-

EDTA (TE) buffer (pH 8.0), 20 mM $MgCl_2$, and 0.05% w/v Tween 20. These relatively low reactant concentrations allowed the reaction to occur at a slow rate while generating a sufficient fluorescence signal for detection. The released output reacts with the reporter through TMSD to separate the fluorophore from the quencher, restoring the fluorescent signal. This reaction also prevents the reverse reaction as the output DNA is consumed by this fast reaction. The reaction progress was monitored in real-time using a fluorescence plate reader, with fluorescence detected every 3 minutes for one hour at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

Because the rate of the uncatalyzed reaction is slow, 100 nM substrate and 100 nM fuel were used for determining k_0. Fluorescence was detected every 5 minutes for two hours. The reaction curves were then used to calculate 1/[S]. k_0 can be obtained by plotting 1/[S] against t because:

$$\frac{1}{[S]} = \frac{1}{[S]_0} + k_0 t \qquad (13)$$

4.2.5 Performing modified nucleotide-assisted DNA catalytic reaction.

Unless stated otherwise, the modified nucleotide-assisted DNA catalytic reaction was conducted using 100μL reaction solutions containing 25 nM substrate, 100 nM fuel, 20 nM reporter, and the desired concentration of catalyst in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. The reaction progress was monitored in real-time using a fluorescence plate reader, with fluorescence detected every 3 minutes for one hour at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

4.2.6 microRNA detection

The reaction mixture was prepared containing 2 nM the translator, 1 nM the mediator, and varying concentrations of microRNA (miRNA) in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. After a 30-minute incubation at room temperature, a 50 µL aliquot of this mixture was transferred into a 96-well plate, with each well containing 50 µL of the reaction solution consisting of 50 nM substrate, 200 nM fuel, and 40 nM reporter in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. The plate was loaded onto a plate reader, and the fluorescence was monitored every 3 minutes for one hour at an excitation wavelength of 488 nm and a detection wavelength of 515 nm.

4.3 Results and discussion

4.3.1 Working principle of modified nucleotide-assisted DNA catalytic reaction.

The modified nucleotide-assisted DNA catalytic reaction uses the incorporation of nucleotide modifications into DNA to destabilize the substrate. These modifications can alter base-pairing and base-stacking of dsDNA, therefore, decreasing the stability of the substrate and increasing the free energy of the substrate. The modified nucleotide-assisted DNA catalytic reaction has a similar working principle as mismatch-aided DNA catalytic reaction except that the mismatches are replaced by modified nucleotides. Specifically, the substate (**S**) is formed by hybridization of domains **2**, **3**, **4** of the oligo **D** with the domains **2***, **3***, **4*** of the output DNA (**O**) (**Figure 4.1 A**). The oligo **D** has an unhybridized domain **1** at its 5'-end that acts as the toehold,

and the output DNA also has an unhybridized domain **5** at its 5'-end, which is used to react with the reporter after the output DNA is dissociated. The modified base is incorporated into the domain **3*** to destabilize the substrate. The fuel (**F**) shares the same domains **2***, **3***, and **4*** with the output DNA but does not have modified base. When the catalyst (**C**) is absent, the reaction rate of the fuel with the substrate is slow because there is no toehold for facilitating this reaction. In the presence of the catalyst, the catalyst facilitates the reaction between the substrate and fuel through a multistep process involving two toehold exchange reactions (**Figure 4.1 B**). Domain **1*** of the catalyst first hybridizes to toehold **1** of the substrate, initiating the first toehold exchange reaction. This branch migration leads to the formation of intermediate ***I*** and the release of the output DNA, exposing the previously blocked toehold **4**. The released output DNA then reacts with the reporter through a toehold-mediated strand displacement, restoring the fluorescence of a previously quenched fluorophore molecule. Subsequently, domain **4*** of the fuel binds to toehold **4**, initiating the second toehold exchange reaction. This reaction then forms the product DNA and regenerates the catalyst. Consequently, the catalyst reacts with another substrate molecule, and the process repeats, leading to linear signal amplification of the catalyst.

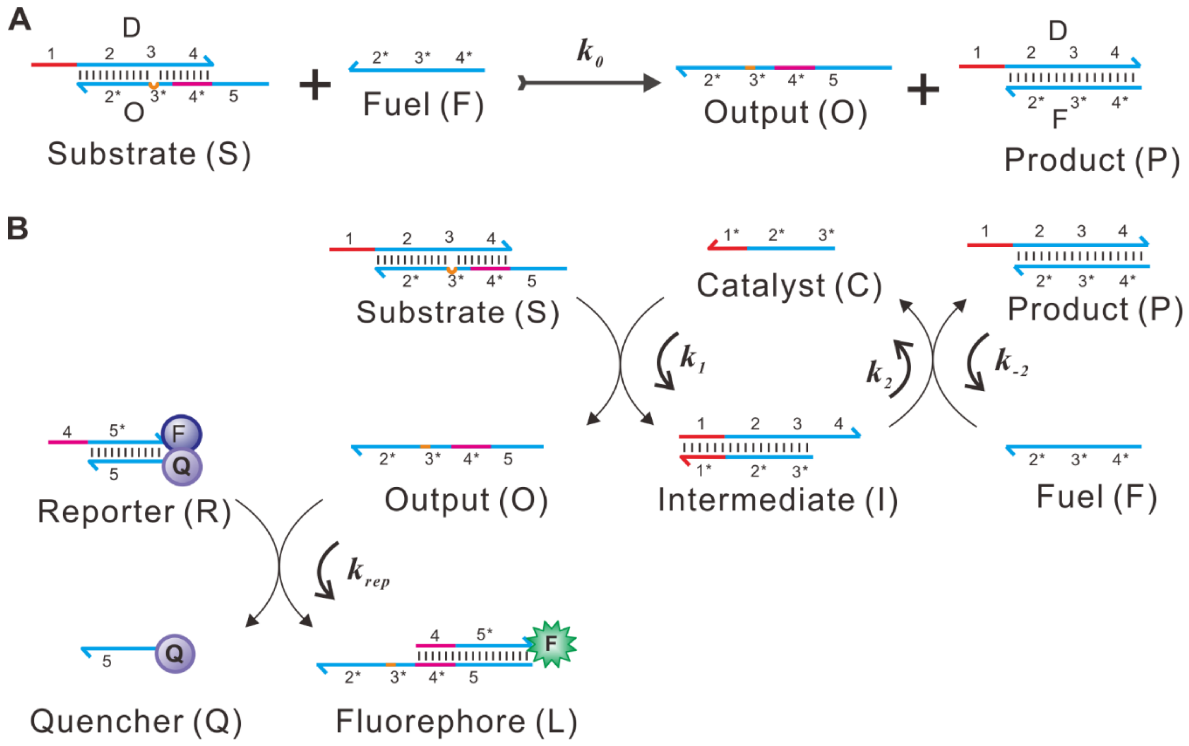


Figure 4.1 *Schemes of modified nucleotide-assisted DNA catalytic reaction.*

(A) Uncatalyzed reaction generating background. (B) Pathway of the catalytic reaction.

Similarly, a two-step model was used to describe the modified nucleotide-aided DNA catalytic reaction, where two toehold exchange reactions are used as the elementary steps:

$$\text{Overall Reaction (OR): } S + F \underset{k_{-1}k_{-2}}{\overset{k_1k_2}{\rightleftarrows}} P + O$$

$$\text{Elementary reaction 1(es1): } S + C \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} I + O$$

$$\text{Elementary reaction 2(es2): } I + F \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} P + C$$

$$\text{Uncatalyzed reaction (uc): } S + F \underset{k_0}{\overset{k_0}{\rightleftarrows}} P + O$$

Signal generation: $O + R \xrightarrow{k_{rep}} E + Q$

If there is no modified nucleotide, the DNA reactants and products are identical, and the free energy change of the overall reaction is zero. The free energy of step 1 depends on the binding energy difference between toeholds **1** and **4**. Since step 2 is the reverse of step 1, its free energy depends on the binding energy difference between toeholds **4** and **1**. In our design, toehold **4** has a lower binding energy than toehold **1**. Thus, the free energy of step 1 is positive and nonspontaneous, while step 2 is negative and spontaneous. Therefore, step 1 is the rate-limiting step. Incorporation of the nucleotide modification into the substrate makes step 1 spontaneous. Therefore, both steps are spontaneous, improving the reaction kinetics.

4.3.2 Designing modified substrates with varying free energy of formations.

I designed five substrates by incorporating three nucleotide modifications, AB, AN, and PS, into the same site in domain **3*** of the output DNA (**Table 4.1**). Modifications AB and AN were incorporated into the output DNA through two approaches: inter-nucleotide insertion (ABi and ANi) and nucleotide replacement (ABr and ANr) (**Figure 4.2 A**). By placing the modified nucleotide in the same locations and relatively close to the center of the substrate, I minimized the effect of location on the kinetics and amplification of the catalytic reaction.

Since no algorithm is available for determining the binding energy of modified nucleotide-containing DNA, I experimentally determined the free energy of formation for these five substrates as described in section 4.2.3, by determining the melting temperature of the dsDNA at varies concentrations using SYBR green (**Figure 4.3**). The free energy of formation of the unmodified substrate, intermediate and products was also experimentally determined for accurate comparison. Interestingly, free energy of the unmodified substrate is slightly lower than that the product

although the substate and product contains the same dsDNA region, which is attributed to the overhang in the substrate [17]. The five modifications, ABi, ANi, ABr, ANr, and PS, resulted in different increases in the free energy of formation of the substrate (ΔG^o_{fs}) (**Table 4.1**). Compared to the ΔG^o_{fs} of the unmodified substrate (-39.5 kcal/mol), the PS modification increased the ΔG^o_{fs} by only 0.6 kcal/mol. The ABi and ANi modifications led to similar ΔG^o_{fs} (-38.1 kcal/mol and -38.2 kcal/mol), increasing the ΔG^o_{fs} by 1.4 kcal/mol and 1.3 kcal/mol, respectively. The ABr and ANr modifications also resulted in similar ΔG^o_{fs} (-35.5 kcal/mol and -35.7 kcal/mol) with larger increases. Therefore, the pairs of modifications, ABi and ANi, and ABr and ANr, are particularly useful for studying the impact of the modification structure on the kinetics and amplification efficiency of catalytic reactions.

The PS modification caused the smallest increase in free energy, suggesting it only slightly affects base stacking at the modification site. However, PS might be advantageous for fine-tuning free energy. Inter-nucleotide insertion has a smaller impact on free energy than nucleotide replacement for AB and AN modifications, as inter-nucleotide insertion only disrupts base-stacking and not base-pairing. For both inter-nucleotide insertion and nucleotide replacement, the AB and AN modifications lead to similar changes in ΔG^o_{fs}, suggesting that despite their different molecular structures, AB and AN have similar impacts on base pairing and base stacking.

Table 4.1 *Sequences used for creating substrates containing modification with varying* ΔG^{o}_{fs}.

Nucleotide modifications are highlighted in ***orange*** *and* ***bolded****.*

	Sequence	ΔG^{o}_{fs} of **D-1:O** (kcal/mol)
D-1	CAT CAC CTC ACT CAA C CTA CCA C ATC TCC AC 1 2 3 4	
O-wt	AGA AGG TGA GTA GTG TGG AGA T GTG GTA G GTT GAG TGA G 5 4* 3* 2*	**-39.5**
O-ABi	AGA AGG TGA GTA GTG TGG AGA T GTG **/idSp/** GTA G GTT GAG TGA G 5 4* 3* 2*	**-38.1**
O-ANi	AGA AGG TGA GTA GTG TGG AGA T GTG **/iSpC3/** GTA G GTT GAG TGA G 5 4* 3* 2*	**-38.3**
O-ABr	AGA AGG TGA GTA GTG TGG AGA T GTG **/idSp/** TAG GTT GAG TGA G 5 4* 3* 2*	**-35.5**
O-ANr	AGA AGG TGA GTA GTG TGG AGA T GTG **/iSpC3/** TAG GTT GAG TGA G 5 4* 3* 2*	**-35.7**
O-PS	AGA AGG TGA GTA GTG TGG AGA T GTG **G*** TA G GTT GAG TGA G 5 4* 3* 2*	**-38.9**
F	G TGG AGA T GTG GTA G GTT GAG TGA G 4* 3* 2*	**-38.6**
Fs	TGG AGA T GTG GTA G GTT GAG TGA G 4* 3* 2*	**-36.9**
C	GTG GTA G GTT GAG TGA G GTG ATG 3* 2* 1*	**-36.1**
Q	5IABkFQ/ -AGA AGG TGA GTA GTG 5	
L	ATC TCC A CA CTA CTC ACC TTC T/FAM/ 4 5*	

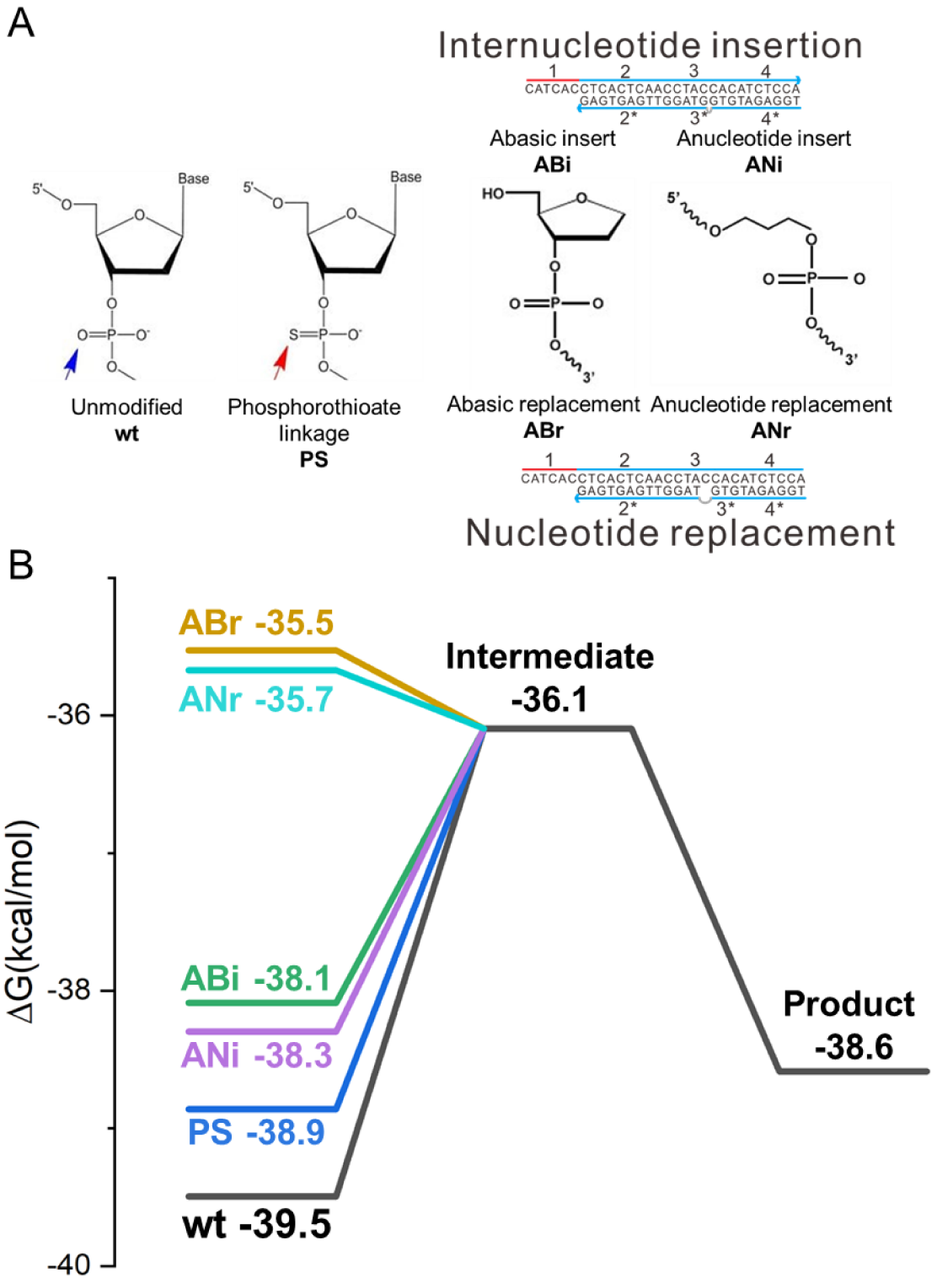


Figure 4.2 *Design of the substrates using different nucleotide modifications (A) and the free-energy change diagram of all the substrate (B).Structure obtained online from Sigma Aldrich and Biosynthesis.*

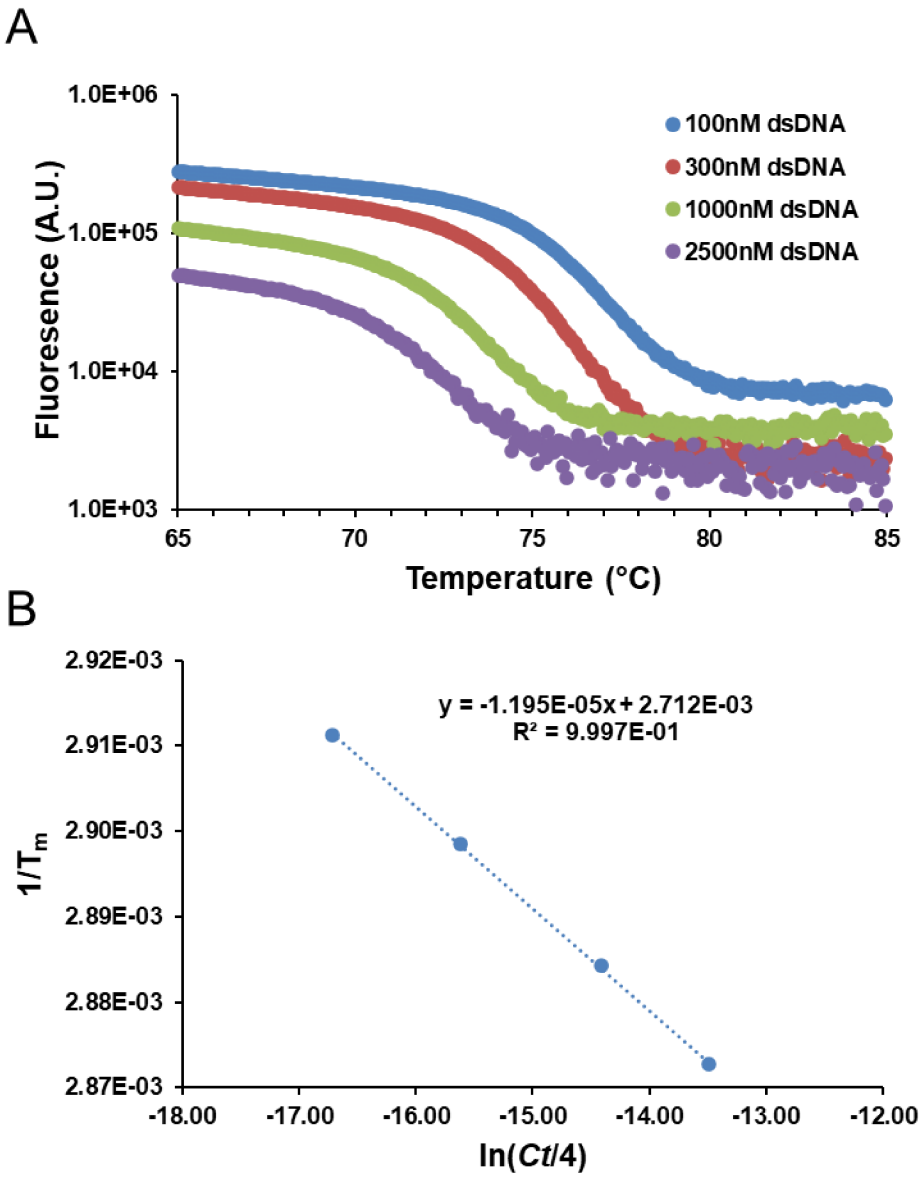


Figure 4.3 *Experimental determination of free energy of the product.*

(A) Melting curves determined using SYBR green, (B) $1/T_m$ plotted against Ln(Ct/4) for determination of ΔH and ΔS using slope and y-intercept.

4.3.3 Kinetics of catalytic reactions using different modified substrates.

I then determined the kinetics of the catalytic reactions using these five modified substrates. This was named modified nucleotide-assisted DNA catalytic reaction (MONA). Two fuels were

used, one with a 7nt-toehold binding domain (TBD) and the other with an 8nt-TBD, to react with these five substrates. Therefore, there were 10 catalytic reactions in total, divided into two groups. The first group involves the reactions of the 7nt-TBD fuel with the five substrates, while the second group involves the reactions of the 8nt-TBD fuel with the five substrates. The two groups share the same five reactions of the elementary step 1, while each group has an identical elementary step 2.

I determined the rate constant k_1 of five step 1 reactions and rate constant k_2 of two step 2 reactions (**Figure 4.4**). Similar to the results observed for the mismatch-aided DNA catalytic reaction, increasing the free energy of the substrate resulted in a continuous increase of the rate constant k_1. The unmodified substrate with the lowest free energy of formation (ΔG^o_{fs}=-38.9 kcal/mol) led to the lowest k_1 (1.56×10^5 $M^{-1}s^{-1}$), while the ABr substrate with highest ΔG^o_{fs} (-35.5 kcal/mol) produced the highest k_1 (2.39×10^5 $M^{-1}s^{-1}$), representing about a 15.3 times improvement in k_1. Interestingly, the two pairs of the substrates, ABi and ANi substrates, and ABr and ANr substrates, have similar ΔG^o_{fs} within each pair, but the k_1 values of the AB substrates are significantly higher than those of the corresponding AN substrates (Figure 4.4). This suggests that the modification structure also significantly impacts the kinetics of the catalytic reactions.

The rate constant k_2 of the elementary step 2 was 1.2×10^6 $M^{-1}s^{-1}$ and 2.7×10^6 $M^{-1}s^{-1}$ for the 7nt-TBD and 8nt-TBD fuels, respectively, which is consistent with the TBD binding energy of two fuels.

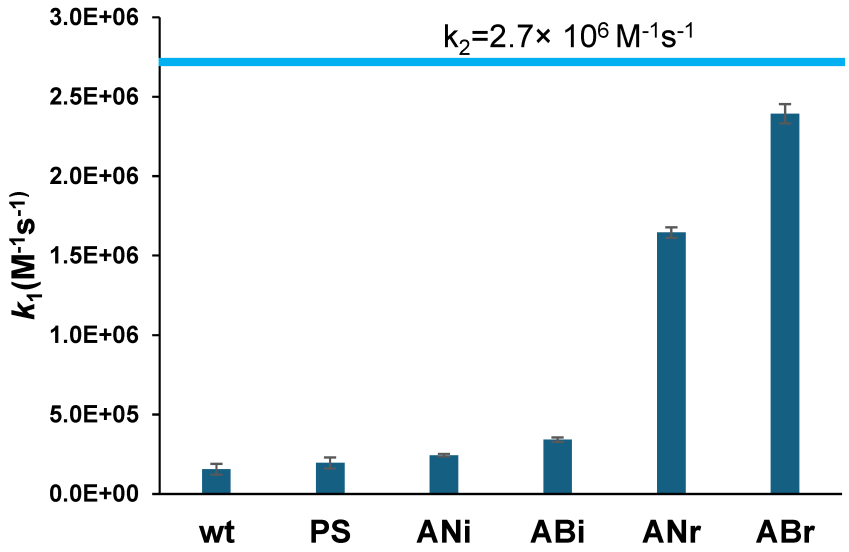


Figure 4.4 *Kinetics of the catalytic reactions using substrates with different nucleotide modifications.*

4.3.4 Kinetics of uncatalyzed reactions using different modified substrates

I then determined the rate constant k_0 of uncatalyzed reactions using different modified substrates. These uncatalyzed reactions, also known as leakages, generate the background in amplification. Although increasing the free energy of formation of the substrate (ΔG^o_{fs}) also increases k_0 values, this increase is much smaller than that in k_1 (**Figure 4.5 A**). For instance, when the 7nt-TBD fuel was used, increasing the ΔG^o_{fs} resulted in a continuous increase of the rate constant k_0. The substrate with the highest ΔG^o_{fs} (-35.5 kcal/mol) had a k_0 of 12.5 $M^{-1}s^{-1}$, which is about 2.3 times lower than the lowest k_0 (28.6 $M^{-1}s^{-1}$) generated by the unmodified substrate with the lowest ΔG^o_{fs} (-38.9 kcal/mol).

Additionally, the 8nt-TBD fuel resulted in higher k_0 values than the 7nt-TBD fuel for all uncatalyzed reactions. This is because the free energy (-38.6 kcal/mol) of the product generated by the 8nt-TBD fuel is lower than that (-36.9kcal/mol) of the product generated by the 7nt-TBD

fuel. The 8nt-TBD fuel increased the k_0 by about 42-104% for five substrates, with the lowest increase for the ANi substrate and highest increase for the PS substrate (**Figure 4.5 A**).

I further calculated k_1/k_0 values for all these reactions. The reaction using the ANr substrate and the 8nt-TBD fuel resulted in the highest k_1/k_0 value (**Figure 4.5 B**). Because the k_1/k_0 value is associated with the signal-to-background ratio, I selected the ANr substrate to perform the modified nucleotide-aided DNA catalytic reaction for testing the amplification efficiency.

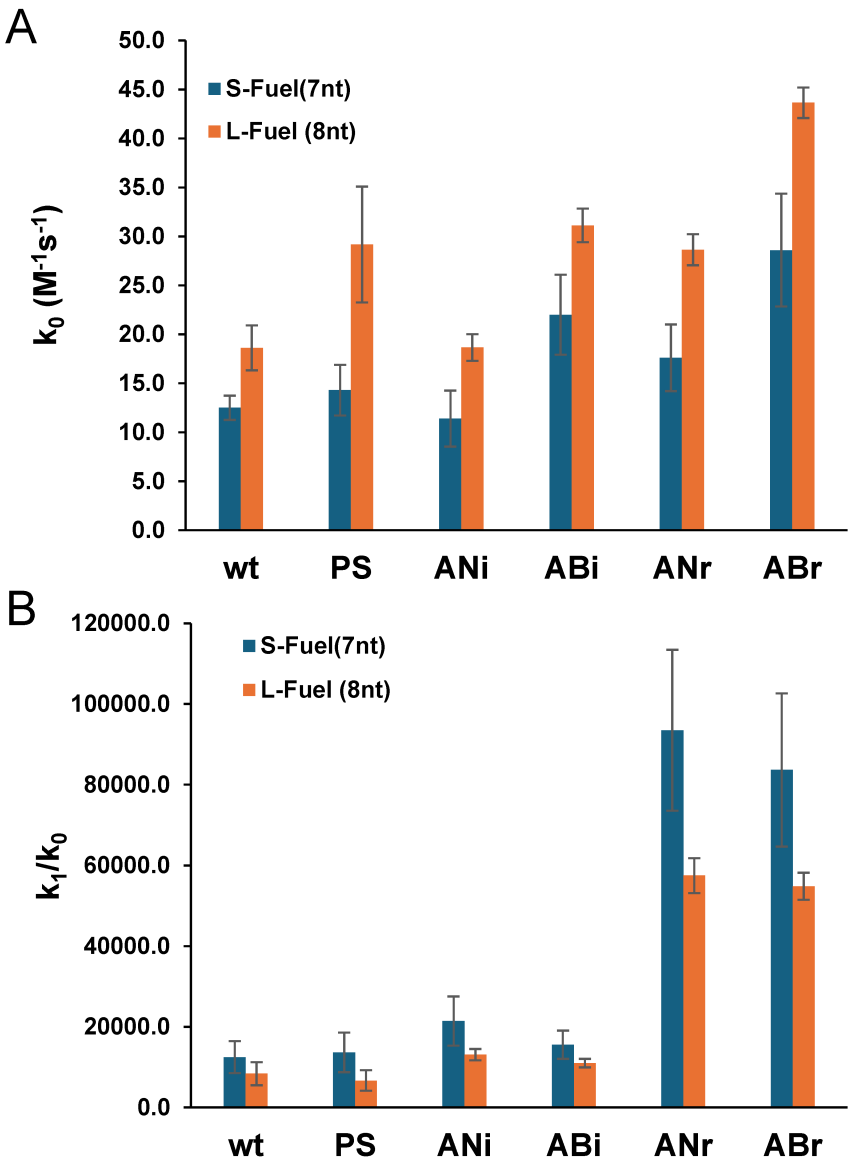


Figure 4.5 *(A) Rate constant k_0 of the uncatalyzed reactions using substrates with different nucleotide modifications; (B) k_1/k_0 of these reactions using substrates with different nucleotide modifications.*

4.3.5 Effect of the ANr location on amplification efficiency of the catalytic reaction.

I further investigated the effect of the ANr location on amplification efficiency of the catalytic reaction. I designed thirteen output (**O**) strands each containing an ANr modification at

different locations (**Table 4.2**). An ANr-incorporated output is denoted as ANr (m, n), Where m indicates the number of nucleotides between the modification and the 3'-end of domain **2*** and n indicates the number of nucleotides between the modification and the 5'-end of domain **4* (Figure 4.1 and Figure 4.6 A).** Using these output strands, I prepared thirteen substrates, each with a ANr modification at different locations. I then performed the modified nucleotide-assisted DNA catalytic reactions with these substrates using 100 pM the catalyst. The signal and background of these reactions were measured after a 1-hour reaction (**Figure 4.6 B**).

Notably, the impact of the ANr location on the background is stronger than on the signal. Substrates with the ANr location proximal to the 3'-end of domain **2*** and 5'-end of domain **4*** resulted in a larger background. In contrast, the background decreased as the ANr location moved towards the middle of the DNA hybrid between domains **2***, **3***, and **4*** with their complements. The substrate with an ANr modification at the location (m, n) = (12, 12) resulted in the lowest background, which is about 10-times lower than the highest background generated by the substrate with an ANr (6, 18). The high background of ANr (6, 18) is likely due to increased DNA end fraying caused by ANr site located near the 3'-end of **2*** through disruption of base pairing and base stacking[7-8]. Therefore, the 6-nucleotide region 3' of the ANr site, or the m arm, on the output strand is likely to dissociate from the **D-1** strand, exposing **D-1** strand and causing high background despite the n arm being stable. Although the outputs ANr (6, 18) and ANr (18, 6) have the ANr site positioned 6-nucleotide away from the ends of dsDNA region, there is a significant lower background for ANr (18, 6) due to the higher GC content at the 5' of **4*** compared to the 3' of **2***.

The ANr location exhibited a weaker impact on the signal than on background. When the results of the substrate with the highest background were excluded, the highest signal was only about 1.4 times higher than the lowest. There is no clear correlation between the ANr location and

the signal. The best signal to background ratio was observed for the substrate with an ANr (12, 12). Additionally, this substrate also led to the highest amplification efficiency. These results are similar to those observed for the TWJ-mediated DNA catalytic reaction.

Table 4.2 *Sequences of the output DNA strands incorporated with a ANr modification at different locations*

	Sequence
O-ANr (6,18)	AGA AGG TGA GTA GTG TGG AGA T GTG GTA G GTT /iSpC3/ AGT GAG 5 4* 3* 2*
O-ANr (7,17)	AGA AGG TGA GTA GTG TGG AGA T GTG GTA G GT /iSpC3/ GAG TGA G 5 4* 3* 2*
O-ANr (8,16)	AGA AGG TGA GTA GTG TGG AGA T GTG GTA G G /iSpC3/ TGA GTG AG 5 4* 3* 2*
O-ANr (9,15)	AGA AGG TGA GTA GTG TGG AGA T GTG GTA G /iSpC3/ TTG AGT GAG 5 4* 3* 2*
O-ANr (10,14)	AGA AGG TGA GTA GTG TGG AGA T GTG GTA /iSpC3/ GTT GAG TGA G 5 4* 3* 2*
O-ANr (11,13)	AGA AGG TGA GTA GTG TGG AGA T GTG GT /iSpC3/ G GTT GAG TGA G 5 4* 3* 2*
O-ANr (12,12)	AGA AGG TGA GTA GTG TGG AGA T GTG G /iSpC3/ AG GTT GAG TGA G 5 4* 3* 2*
O-ANr (13,11)	AGA AGG TGA GTA GTG TGG AGA T GTG /iSpC3/ TAG GTT GAG TGA G 5 4* 3* 2*
O-ANr (14,10)	AGA AGG TGA GTA GTG TGG AGA T GT /iSpC3/ GTA G GTT GAG TGA G 5 4* 3* 2*
O-ANr (15,9)	AGA AGG TGA GTA GTG TGG AGA T G /iSpC3/ GGT AG GTT GAG TGA G 5 4* 3* 2*
O-ANr (16,8)	AGA AGG TGA GTA GTG TGG AGA T /iSpC3/ T GGT AG GTT GAG TGA G 5 4* 3* 2*
O-ANr (17,7)	AGA AGG TGA GTA GTG TGG AGA /iSpC3/ GTG GTA G GTT GAG TGA G 5 4* 3* 2*
O-ANr (18,6)	AGA AGG TGA GTA GTG TGG AG /iSpC3/ T GTG GTA G GTT GAG TGA G 5 4* 3* 2*

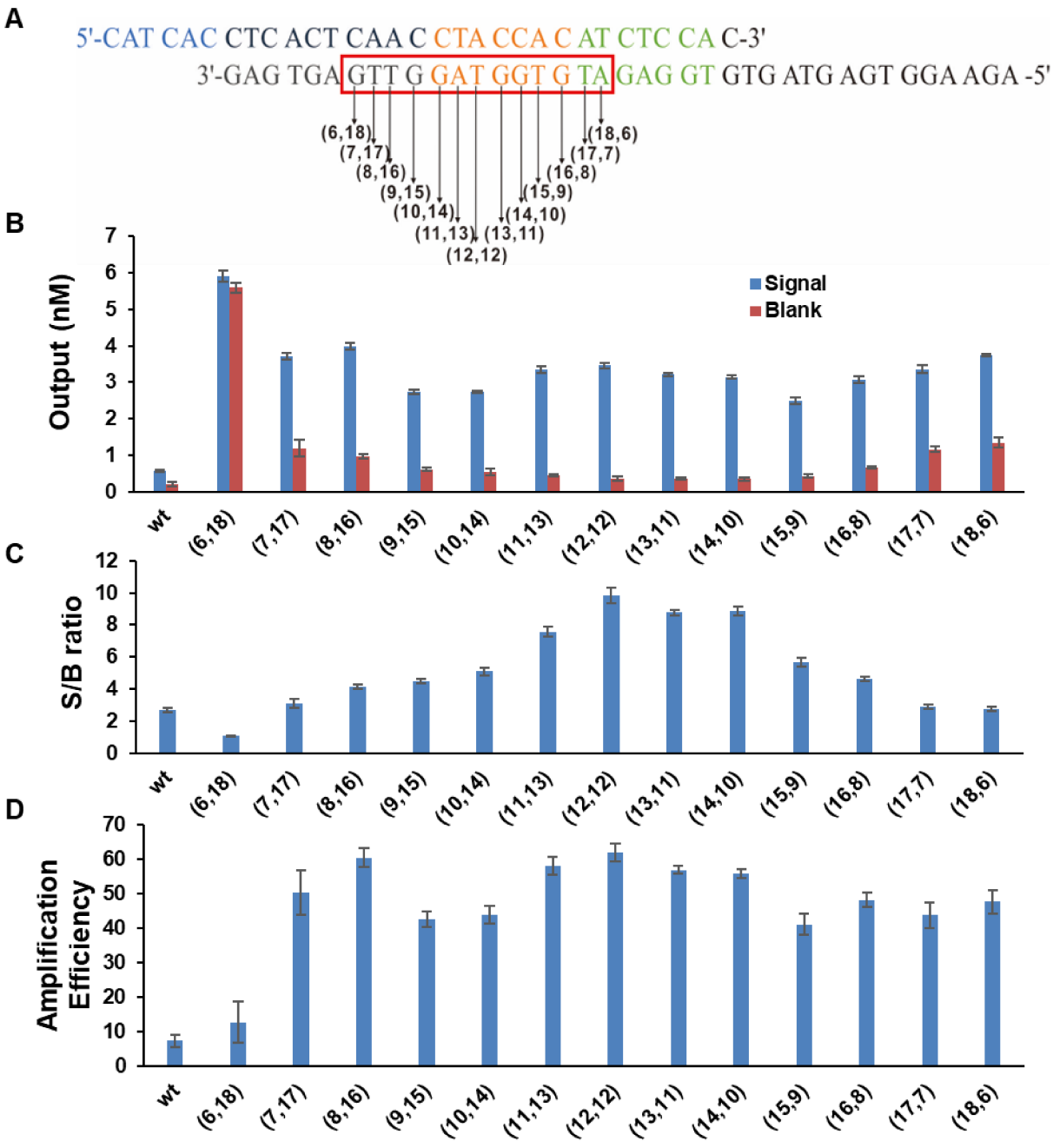


Figure 4.6 *Effect of the ANr location on amplification efficiency of the catalytic reaction.*

(A) Locations of the ANr modification; (B) The signal amplification and background of reaction (C) Signal to background ratio of each substrate (D) the amplification efficiency of each substrate

4.3.6 Using modified nucleotide-assisted DNA catalytic reaction as an assay for amplified detection of the catalyst.

Modified nucleotide-assisted DNA catalytic reaction (MONA) can be directly used as an assay for the amplified detection of the catalyst. To achieve high detection sensitivity, I used the substrate containing the ANr at the (12, 12) location for performing the catalytic reaction. I also optimized other conditions and the optimized conditions are 25 nM substrate, 100 nM fuel, 20 nM reporter in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20.

I then used the optimized conditions to determine the dynamic range and LOD of the catalytic reaction as an assay (**Figure 4.7**). A linear dynamic range from 5 pM to 500 pM was obtained. The assay was able to detect as little as 2.4 pM of the catalyst within 1 hour at room temperature without the need for any enzymes. Therefore, MONA shows the similar dynamic range and LOD shows the similar dynamic range and LOD to the TWJ-mediated DNA catalytic reaction and the mismatch-assisted DNA catalytic reaction.

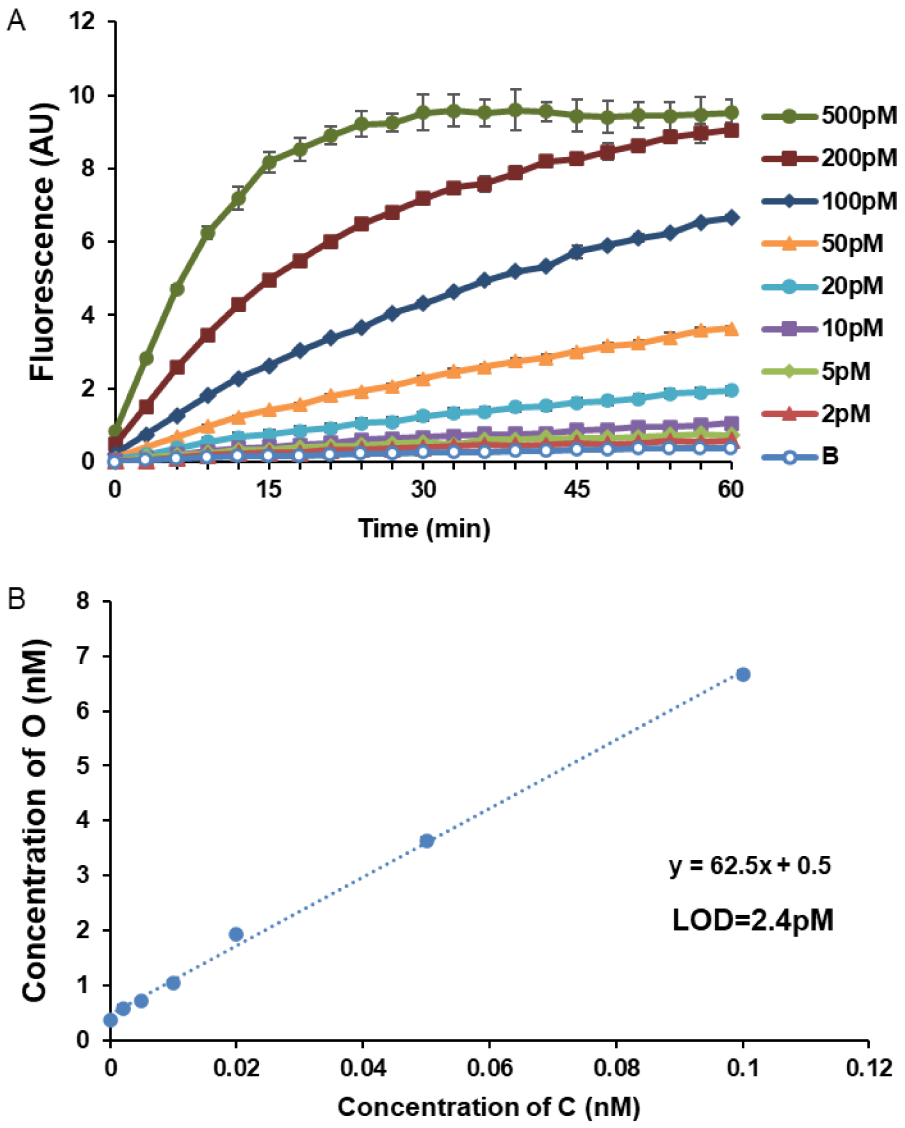


Figure 4.7 *Dynamic range and limit of detection of the modified nucleotide-aided DNA catalytic reaction. A) Fluorescence amplification curves B) Dynamic range, limit of detection, and amplification efficiency. Amplification efficiency is the slope of the linear range.*

4.3.7 Developing a translator and coupling it to the catalytic reaction for detection of miRNA

To apply the modified nucleotide-aided DNA catalytic reaction to the detection of other nucleic acids, such as miRNA[12-16], I designed a DNA translator that converts a miRNA target into a catalyst to initiate the catalytic reaction. Using the translator, I can detect different miRNA targets with the same catalytic reaction by simply changing the design of the translator. The main advantage of using the translator is that high amplification efficiency can be achieved for various nucleic acid targets, even those with stable secondary structures. The stable secondary structure, particularly involving the toehold binding domain, can significantly reduce the kinetics of the toehold exchange reaction and thereby the amplification efficiency of the catalytic reaction. However, since the translator involves a single step of toehold-mediated strand displacement reaction, the stable secondary structure of the nucleic target has a less impact on the overall reaction kinetics.

I designed the translator system to comprises two components, a translator formed by hybridization of the strand **P** with the catalyst, and a single-stranded mediator strand (**M**) (**Table 4.3, Figure 4.8**). The strand **P** includes a toehold (**miR*a**) complementary to a portion of the miRNA target. In the absence of the miRNA target, the translator does not react with the mediator because the toehold-binding domain is absent, and the catalyst is trapped in the translator. In the presence of the miRNA target, the miRNA target binds to the mediator, placing the toehold binding domain and displacing domain (**1*** and **2***) together. Consequently, a toehold-mediated strand displacement reaction is initiated, releasing the catalyst to trigger the modified nucleotide-aided DNA catalytic reaction.

I first used a DNA sequence identical to the miRNA target miRNA-10b to develop and optimize the translator. I then coupled the translator to the modified nucleotide-aided DNA catalytic reaction for amplified detection of miRNA 10b. 50 μL of the mixture was prepared containing 2 nM the translator, 1 nM the mediator, and varying concentrations of miRNA 10 (miRNA) in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. After a 30-min incubation at room temperature, the mixture was added to a well preloaded 50 μL of reaction solution containing 50 nM substrate, 200 nM fuel, and 40 nM reporter in 10 mM TE buffer (pH 8.0) with 20 mM $MgCl_2$ and 0.05% w/v Tween 20. The fluorescence of the reaction was then monitored every 3 min for one hour.

I first applied the translator-catalytic reaction system to detect miRNA 10b in buffer, achieving a dynamic range of 10 pM to 200 pM and LOD of 6.5 pM (**Figure 4.9**). Next, I used the system to detect miRNA 10b spiked in cell lysate, prepared by lysing 2 million MB 231 cells into 2 mL of lysis buffer. A buffer control and a non-spiked lysate control were run to ensure there were no false positive signals. The spiked amount mimicked intracellular conditions as the intracellular miRNA 10b is estimated to be 600 copies or 540-630 pM[18], based on a cell diameter of 15μm[19,20]. A similar dynamic range (10 pM to 200 pM) and LOD (8.2 pM) were obtained (**Figure 4.10**), demonstrating the applicability of the system for detecting miRNA 10b in biological samples.

Table 4.3 Sequences of the miRNA target and the translator for miRNA 10b	
	Sequence
miRNA 10b	UAC CCU GUA GA ACC GAA UUU GUG a b
miRNA 10b DNA	TAC CCT GTA GA ACC GAA TTT GTG a b
P mir10b	CAC AAA TTC GGT CAT CAC CTC ACT CAA C miR b* 1 2
M mir10b	GTT GAG TGA G GTG ATG ta TCT ACA GGG TA 2* 1* miR a*
Blocker	GTT GAG TGA G GTG ATG 2* 1*

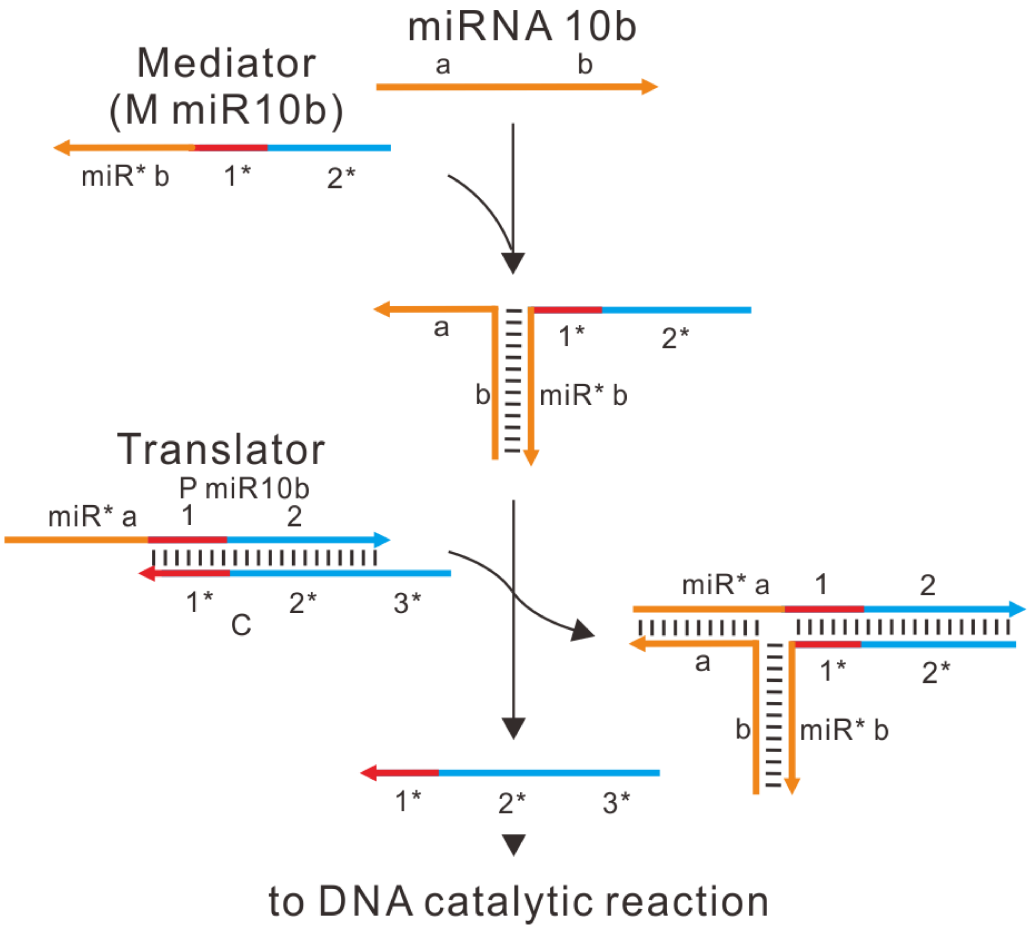


Figure 4.8 *Design and schematics of miRNA translator.*

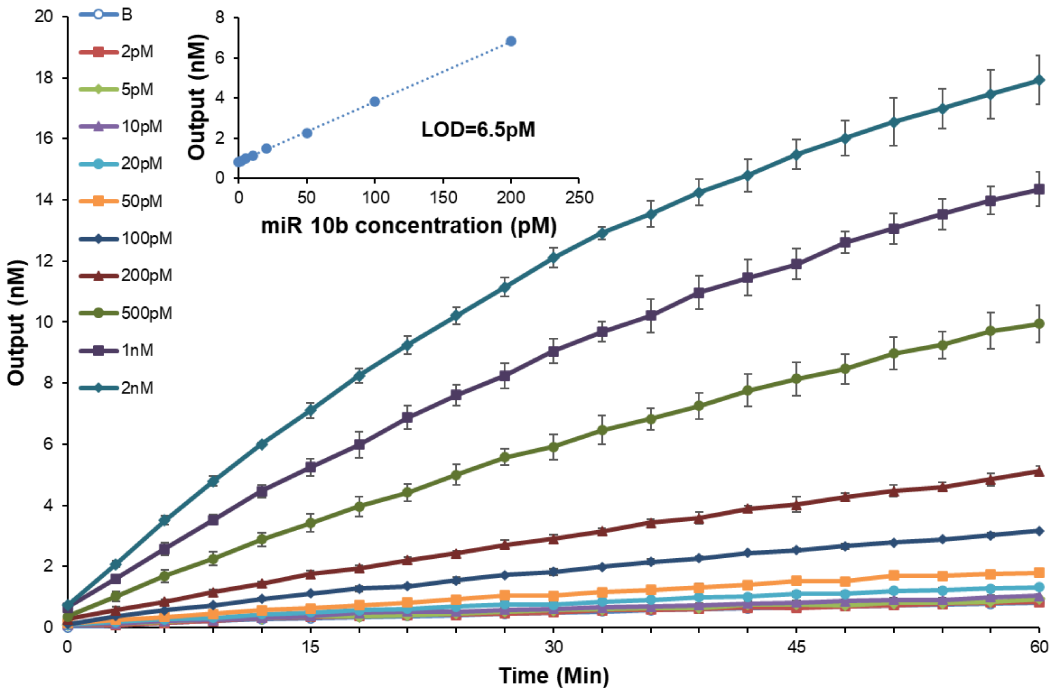


Figure 4.9 *Dynamic range and limit of detection of the translator-catalytic reaction system for detection of miRNA 10b in buffer.*

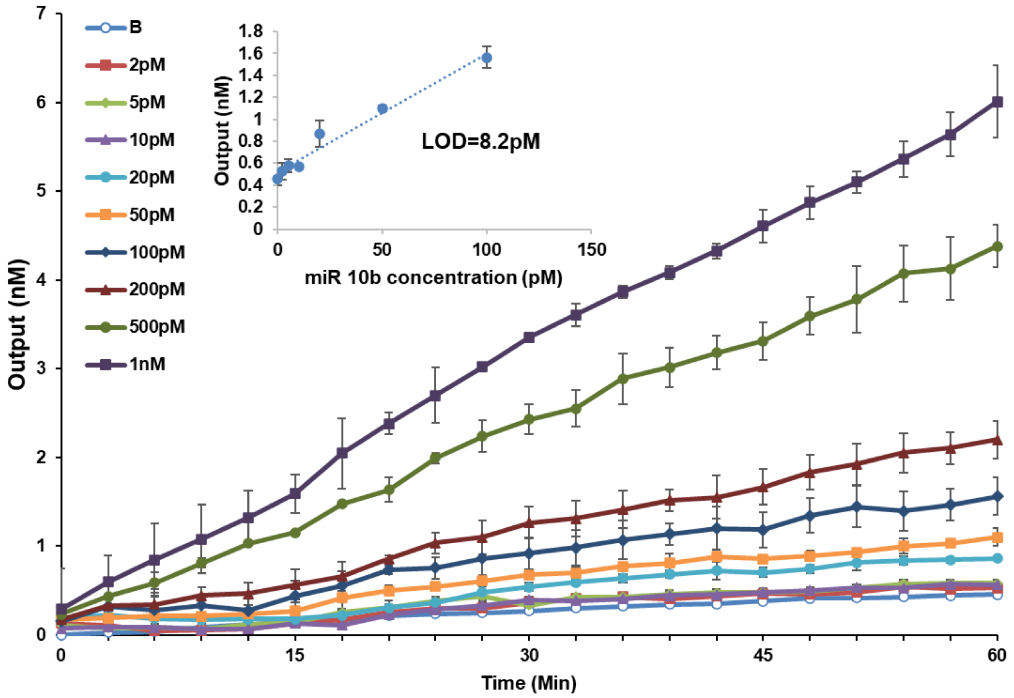


Figure 4.10 *Dynamic range and limit of detection of the translator-catalytic reaction system for detection of miRNA 10b in cell lysate.*

4.4 Conclusion

I exploited nucleotide modifications to enhance the kinetics of toehold exchange catalytic reactions and developed a modified nucleotide-assisted DNA catalytic reaction. Five substrates with different nucleotide modifications, ABi, ANi, ABr, ANr, and PS, were designed and prepared. Their free energies were experimentally determined, showing varying increases. Compared to the unmodified substrate, the PS modification increased the free energy by only 0.6 kcal/mol. The ABi and ANi modifications resulted in similar free energies, increasing by 1.4 kcal/mol and 1.3 kcal/mol, respectively. The ABr and ANr modifications also led to similar free energies with larger increases.

Kinetic studies revealed that increasing the substrate's free energy resulted in a continuous increase in the rate constant k_1 of elementary step 1, while its impact on the rate constant k_0 of leakages was less significant. The ANr-modified substrate yielded the highest k_1/k_0 value, indicative of the optimal signal-to-background ratio in the catalytic reaction.

The ANr substrate was selected for further testing of amplification efficiency. The effect of the ANr location on amplification efficiency was also investigated. The location of ANr had a stronger impact on the background than on the signal. The substrate with an ANr (12, 12) modification exhibited the lowest background, approximately 10 times lower than the highest background generated by the substrate with an ANr (6, 18). Conversely, the highest signal was only about 1.4 times higher than the lowest among thirteen ANr locations. The ANr (12, 12) location achieved the highest amplification efficiency, at 60-fold per hour.

As an assay for catalyst detection, the catalytic reaction could detect as low as 2.4 pM catalyst at room temperature without requiring enzymes. A translator was developed to enable the detection of different nucleic acid targets using the same catalytic reaction. This ensures that the

secondary structures of the nucleic acid targets do not affect the amplification efficiency of the catalytic reaction. The translator, coupled with the catalytic reaction, facilitated the amplified detection of miRNA 10b in cell lysate, achieving a LOD of 8.2 pM.

4.5 References

1. Duffy, K., Arangundy-Franklin, S., & Holliger, P. (2020). Modified nucleic acids: replication, evolution, and next-generation therapeutics. BMC Biology. 18, 1122.
2. Jafari, D., Shajari, S., & Samadikuchaksaraei, A. (2020). Modified nucleic acids: Expanding the capabilities of functional biomaterials. MDPI Molecules. 25(20), 46594.
3. Caruthers, M.H. (1985). Chemical synthesis of DNA and RNA: An overview. Bioorg Chem. 13(3), 217.
4. Hoose, A., Vellacott, R., Storch, M., Freemont, P.S., & Ryadnov, M.G. (2023). DNA synthesis technologies to close the gene writing gap. Nat Rev Chem. 7, 144.
5. Beaucage, S. L., & Caruthers, M. H. (1981). Solid-phase synthesis of oligodeoxyribonucleotides by the phosphoramidite method. Tetrahedron Lett. 22(20), 1855.
6. Kabza, A. M., Kundu, N., Zhong, W., & Sczepanski, J. T. (2021). Integration of chemically modified nucleotides with DNA strand displacement reactions for applications in living systems. WIREs Nanomedicine and Nanobiotechnology. 14(2), e1700.
7. Vesnaver, G., Chang, C. N., Eisenberg, M., Grollman, A. P., & Breslauer, K. J. (1989). Influence of abasic and anucleosidic sites on the stability, conformation, and melting behavior of a DNA duplex: Correlations of thermodynamic and structural data. Biochem. 86, 3614.
8. Gelfand, C. A., Plum, G. E., Grollman, A. P., Johnson, F., & Breslauer, K. J. (1998). Thermodynamic consequences of an abasic lesion in duplex DNA are strongly dependent on base sequence. Biochem. 37(20), 7321.
9. Ririe, K. M., Rasmussen, R. P., & Wittwer, C. T. (1997). Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal Biochem. 245(2), 154.
10. Jacob Souza, K., & Agrawal, D. K. (2024). Employing toehold-mediated DNA strand displacement reactions for biomedical applications. Med-X. 2(1), 13.

11. Peng, H., Newbigging, A. M., Reid, M. S., Uppal, J. S., Xu, J., Zhang, H., & Le, X. C. (2019). Signal Amplification in Living Cells: A Review of microRNA Detection and Imaging. Anal Chem. 91(24), 15247.
12. Mariana, L.Q., Rauhut, R., Lendeckel, W., Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science. 294, 853.
13. Volinia, S., Calin, G. A., Liu, C. G., Ambs, S., Cimmino, A., Petrocca, F., et al. (2006). A MicroRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci USA, 103, 2257.
14. Bartel, D.P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell. 116, 281.
15. Dong, H., Lei, J., Ding, L., Wen, Y., Ju, H., Zhang, X. (2013). MicroRNA: Function, detection, and bioanalysis. Chem Rev. 113, 6207.
16. Condrat, C. E., Thompson, D. C., Barbu, M. G., Bugnar, O. L., Boboc, A., Cretoiu, D., Suciu, N., Cretoiu, S. M., & Voinea, S. C. (2020). miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. Cells. 9(2), 276
17. SantaLucia Jr., J., Hicks, D. (2004). The thermodynamics of DNA structural motifs. Annu Rev Biophys. 33, 415.
18. Ma, L., Teruya-Feldstein, J., & Weinberg, R. A. (2007). Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature. 449(7163), 682.
19. Ren, X., Ghassemi, P., Babahosseini, H., Strobl, J. S., & Agah, M. (2017). Single-Cell Mechanical Characteristics Analyzed by Multiconstriction Microfluidic Channels. ACS Sensors. 2(2), 290.
20. Geiger, F., Ru¨diger, D., Zahler, S., Engelke, H. (2019). Fiber stiffness, pore size and adhesion control migratory phenotype of MDA-MB-231 cells in collagen gels. PLoS ONE. 14(11), e0225215.

5. Chapter 5: Conclusion

5.1 Advancement in knowledge

Non-covalent DNA catalytic reactions represent a class of molecular amplification techniques in the field of dynamic DNA nanotechnology[1-3]. These reactions involve the cyclic interaction of a catalyst DNA with DNA substrates through toehold-mediated strand displacement reactions, resulting in the production of multiple DNA products from a single catalyst DNA strand. The DNA substrates are intentionally designed to have kinetically trapped structures, leading to a high activation energy for their uncatalyzed reactions. Consequently, in the absence of the catalyst, the reaction rate is slow, and the DNA substrates remain stable[1-3]. However, when the catalyst is present, it reacts with the DNA substrate, disrupting its stable secondary structure and lowering the activation energy. As a result, the reaction kinetics improve significantly. Similar to a chemical catalyst, the DNA catalyst actively participates in the reaction but is continuously regenerated, allowing it to facilitate the generation of multiple product molecules from a single DNA catalyst strand.

Since the advent of the first DNA catalytic reaction[3-5], three major types of DNA catalytic reactions have been developed and widely used for detection of nucleic acids and other molecules. These include hybridization chain reaction (HCR)[6-8], entropy-driven catalytic reaction (EDC)[9,10], and catalytic hairpin assembly (CHA)[11, 12]. Nevertheless, these catalytic reactions face limitations that affect their suitability for biosensing applications. For instance, the amplification efficiency of HCR and EDC is typically about 2-5 fold per hour. HCR and CHA uses DNA hairpin as substrates, which can be difficult to design and synthesize. To prepare the high-quality substrates, purification such as polyacrylamide gel electrophoresis (PAGE) is often needed[1-3]. To overcome

these limitations, I utilize the toehold exchange catalytic reaction to develop the catalytic reaction that is simple and has improved amplification efficiency while maintaining low background due to leakage reactions.

5.1.1 Development of three-way-junction mediated DNA catalytic reaction

I developed a simple and fast DNA catalytic reaction, termed the TWJ-mediated catalytic reaction, by introducing a TWJ into the substrate. Without the TWJ, the product of the toehold exchange catalytic reaction is identical to the reactants. Consequently, the free energy change of the toehold exchange catalytic reaction is zero, unfavorable for reaction kinetics, and resulting in low amplification efficiency. Introducing a TWJ into the substrate disrupts the normal base-stacking in dsDNA to destabilizes it, thereby increasing the free energy of formation of the substrate (ΔG_{fs}^{o})[13-15]. Because the product DNA remains unchanged, the free energy change of the overall reaction decreases to about -7 kcal/mol, making the reaction spontaneous.

I have demonstrated that the amplification efficiency of TWJ-mediated DNA catalytic is 35-fold per hour or 14 times higher than the toehold exchange catalytic reaction using substrate without TWJ, which suggest increasing the ΔG_{fs}^{o} can enhance the kinetics of catalytic reactions. Another benefit of the TWJ structure is it only causes a slight increase in background leakage because the base pairs in the original dsDNA are unchanged. I found the background leakage remains low when the binding of the two arms is stable.

The high amplification efficiency and low leakage of the TWJ-mediated catalytic reaction allows for detection of as low as 2 pM the catalyst at room temperature without using any enzymes. To further improve the amplification efficiency, I constructed a two-layer TWJ-mediated catalytic

reaction. This two-layer system exhibited signal amplification of 250-fold in one-hour, 1000-fold in two-hour, and 1950-fold in a three-hour lowering the limit of detection to 0.3 pM for the catalyst.

In addition to its high amplification efficiency, the TWJ substrate is easily prepared by annealing two or three oligonucleotide strands without requiring any purification. The high sensitivity, enzyme-free nature, and ease of use make the TWJ-mediated DNA catalytic reaction promising for *in situ* signal amplification as well as on-site analysis and point-of-care applications.

5.1.2 Development of a mismatch-aided DNA catalytic reaction

Inspired by the findings of our TWJ reaction, I developed a simple and flexible mismatch-aided DNA catalytic reaction that is particularly suitable for studying correlations between free energy of the substrate and rate constants of catalyzed and uncatalyzed reactions. I achieved effective modulation of the free energy of formation of the substrate (ΔG^o_{fs}) , and consequently the free energy changes of elementary steps and overall reaction, by strategically manipulating the location, base, and number of mismatches [15, 16]. I designed 14 substrates with ΔG^o_{fs} ranging from -39.1 kcal/mol to -28.3 kcal/mol to conduct the mismatch-aided DNA catalytic reactions. The rate constant of the elementary step 1, k_1, and rate constant of the uncatalyzed reactions, k_0, were determined for each substrate. The relationship between the ΔG^o_{fs} and the rate constants k_1 *and* k_0 is examined.

Interestingly, a linear relationship was observed between the ΔG^o_{fs} and rate constant k_1. This suggests that increasing the ΔG^o_{fs} proportionally enhances the rate constant k_1. In contrast, a nonlinear relationship was observed between the ΔG^o_{fs} and rate constant k_0. As the ΔG^o_{fs} increases, k_0 increases at an accelerating rate. In addition, increasing the ΔG^o_{fs} from -39.1 kcal/mol to -28.3

kcal/mol impact the k_0 less than the k_1. Increasing the ΔG^o_{fs} by about 7 kcal/mol higher than the free energy of formation of the product DNA has no significant effect the k_1/k_0 ratio.

Using a mismatch substate with a free energy of -34.4 kcal/mol, I further tested the amplification efficiency of the mismatch-aided DNA catalytic reaction. Compared to the toehold exchange catalytic reaction using a substrate without mismatch the mismatch-aided DNA catalytic reaction exhibits a 7.4 times higher amplification efficiency, with an amplification efficiency of 47-fold in one hour. Using the mismatch-aided DNA catalytic reaction as an assay for detection of the catalyst, as low as 2 pM of catalyst can be detected at room temperature without the need for any enzymes. Additionally, I further improved the amplification efficiency by developing a two-layer mismatch-aided catalytic reaction that exhibited a 2060-fold signal amplification in a three-hour reaction. This two-layer reaction achieved a limit of detection of 0.5 pM of catalyst.

The findings of this chapter revealed the relationship between the free energy change and kinetics of the toehold exchange catalytic reactions to provide valuable suggestions for designing new toehold exchange catalytic reactions.

5.1.3 Development of a modified nucleotide-assisted DNA catalytic reaction

I developed a modified nucleotide-aided DNA catalytic reaction using nucleotide modifications[17-18] to enhance the kinetics of toehold exchange catalytic reactions. I used different nucleotide modifications, specifically ABi, ANi, ABr, ANr, and PS, to design 5 substrates, and experimentally determined their free energies of formation (ΔG^o_{fs}). ANi and ABi had similar ΔG^o_{fs} as did ANr and ABr. The increase in ΔG^o_{fs} compared to the unmodified substrate were 0.6 kcal/mol for PS, 1.3 kcal/mol for ANi, 1.4 kcal/mol for ABi, 3.8 kcal/mol for ANr and 4.0 kcal/mol for ABr.

Kinetic studies revealed the increase ΔG^o_{fs} led to increase in the rate constant of elementary step 1, k_1. However, the ΔG^o_{fs} had a less significant impact on the rate constant of leakages, k_0. The k_1/k_0 ratio was the highest for the ANr-modified substrate, indicative of the optimal signal-to-background ratio in the catalytic reaction.

The ANr modification was then selected to determine the effect of the location of the modified nucleotide on amplification efficiency. The location of ANr had a bigger impact on the background than on the amplification signal. Among the different substrate variants, ANr (12, 12) generated the lowest background, nearly ten times lower than the background produced by the ANr (6, 18) substrate, which had the highest background. In contrast, the highest signal was only about 1.4 times higher than the lowest signal among the thirteen ANr locations tested. Notably, the ANr (12, 12) location achieved the highest amplification efficiency, at 62-fold per hour.

Applying the catalytic reaction for the enzyme-free, room temperature detection of catalyst, as low as 2.4 pM catalyst can be detected, with an amplification efficiency of 62-fold per hour. To enable the detection of different nucleic acid targets using the same catalytic reaction without affecting the amplification efficiency, I developed a translator. Coupling the translator with the catalytic reaction facilitated the amplified detection of miRNA 10b in cell lysate, achieving a LOD of 8.2 pM.

5.2 Future research

Future research could apply the developed reactions to intracellular detection of nucleic acid biomarkers for diagnosis of disease[2-3, 19-21]. These techniques could also be applied to studying the dynamic functions and profiles of nucleic acids, such as miRNA in living cells. The DNA

catalytic reaction can operate isothermally at a range of temperatures, including room temperature and 37°C, without the need for any enzymes, which is ideal for intracellular application. Challenges of intracellular detection include the need of transfection agent for cellular uptake and degradation by intra-and extracellular nucleases. To promote cellular uptake, gold nanoparticles and 3-dimensional DNA nanostructure such DNA tetrahedron could be incorporated because they have been shown can promote cellular uptake by mammalian cells without the need of any transfection agents in previous studies[21-23]. However, the denes packing of DNA substrate onto the scaffolding can affect its interaction with either the catalyst or fuel and may require further optimization. Chemically modified nucleotide such as phosphorothioate linkage has also been shown to impart resistant against nuclease degradation[17-18] and can be incorporated to further protect the DNA substrate and fuel against.

In addition to amplified detection of intracellular miRNA, the DNA catalytic reactions can be utilized to achieve targeted drug delivery and improve selective drug targeting and reduce side effects[19-21]. The construction of an intracellular target-initiated DNA nanomachine based on DNA catalytic reaction onto gold nanoparticle scaffolding enables efficient cellular entry via endocytosis, high specificity, and higher local effective concentration for improved sensitivity[24-26]. Endosomal escape can be facilitated by the proton sponge effect or through incorporation of pore forming proteins[27]. Because the nanomachine is initiated by intracellular targets, if it is taken up by the normal cells, the system remains inactive, reducing side effects. Combining ligand-receptor mediated drug delivery and intracellular target-triggered drug release using DNA nanomachines can further improve selective drug targeting and minimize side effect.

Finally, further amplification could be achieved through the construction of exponential amplification, where the released output strand can also serve as the catalyst to react with substrates

and fuels, initiating more catalytic reactions. These reactions could enable the detection of sub pico-molar concentration of target with greater amplification efficiency and turn-over time to enable higher sensitivity using shorter amount of time for point of care applications. However, a major challenge for designing such a reaction is high background or leakage, therefore complex substrate purification and careful design of oligonucleotide sequence is often needed. To avoid complex substrate preparation, careful design of the DNA oligonucleotide sequence is needed as well as incorporation of modified nucleotides such as LNA that stabilizes the substrate DNA to help achieve this goal.

5.3 Reference

1. Zhang, D. Y., Seelig, G., (2011) Dynamic DNA nanotechnology using strand-displacement reactions. Nat Chem. 3, 103.
2. Jung, C., Ellington, A. D. (2014) Diagnostic applications of nucleic acid circuits. Acc Chem Res. 47, 1825
3. Simmel, F. C., Yurke, B., & Singh, H. R. (2019). Principles and applications of nucleic acid strand displacement reactions. Chem Rev. 119(8), 6326.
4. Turberfield, A. J., Mitchell, J. C., Yurke, B., Mills, A. P., Blakey, M. I., & Simmel, F. C. (2003). DNA fuel for free-running nanomachines. Phys Rev Lett. 90(11), 118102.
5. Seelig, G., Yurke, B., & Winfree, E. (2006). Catalyzed relaxation of a metastable DNA fuel. J Am Chem Soc. 128(30), 12211.
6. Dirks, R. M., Pierce, N. A. (2004) Triggered amplification by hybridization chain reaction. Proc Natl Acad Sci USA. 101, 15275.
7. Yin, P., Choi, H. M. T, Calvert, C. R., Pierce, N. A. (2008) Programming biomolecular self-assembly pathways. Nature. 451, 318.

8. Venkataraman, S., Dirks, R. M., Rothemund, P. W. K., Winfree, E., & Pierce, N. A. (2007). An autonomous polymerization motor powered by DNA hybridization. Nat Nanotechnol. 2(8), 490.
9. Zhang, D. Y., Turberfield, A. J., Yurke, B., Winfree, E. (2007). Engineering entropy-driven reactions and networks catalyzed by DNA. Science. 318, 1121.
10. Zhang, D. Y., Winfree, E. (2009). Control of DNA strand displacement kinetics using toehold exchange. J Am Chem Soc. 131, 17303.
11. Jiang, Y. S., Li, B., Milligan, J. N., Bhadra, S., Ellington, A. D. (2013). Real-time detection of isothermal amplification reactions with thermostable catalytic hairpin assembly. J Am Chem Soc. 135, 7430.
12. Chen, X., Briggs, N., McLain, J. R., Ellington, A. D. (2013). Stacking nonenzymatic circuits for high signal gain. Proc Natl Acad Sci USA. 110, 5386.
13. Leontis, N. B., Kwok, W., Newman, J. S. (1991). Stability and structure of three-way DNA junctions containing unpaired nucleotides. Nucleic Acids Res. 19, 759.
14. Carr, C. E., Marky, L. A. (2017) Investigation of the melting behavior of DNA three-way junctions in the closed and open states. Biophys J. 113, 529.
15. SantaLucia Jr. J, Hicks D. (2004) The thermodynamics of DNA structural motifs. Annu Rev Biophys. 33, 415.
16. Rossetti, G., Dans, P. D., Gomez-Pinto, I., Ivani, I., Gonzalez, C., & Orozco, M. (2015). The structural impact of DNA mismatches. Nucleic Acids Res. 43(8), 4309.
17. Duffy, K., Arangundy-Franklin, S., & Holliger, P. (2020). Modified nucleic acids: replication, evolution, and next-generation therapeutics. BMC Biology. 18, 1122.
18. Jafari, D., Shajari, S., & Samadikuchaksaraei, A. (2020). Modified nucleic acids: Expanding the capabilities of functional biomaterials. MDPI Molecules. 25(20), 46594
19. Mariana, L. Q., Rauhut, R., Lendeckel, W., Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. Science. 294, 853.
20. Volinia, S., Calin, G. A., Liu, C. G., Ambs, S., Cimmino, A., Petrocca, F., et al. (2006) A MicroRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci USA. 103, 2257.

21. Liang, L., Li, J., Li, Q., Huang, Q., Shi, J., Yan, H., et al. (2014). Single-particle tracking and modulation of cell entry pathways of a tetrahedral DNA nanostructure in live cells. Angew Chem Int Ed Engl. 53, 7745
22. Wang, P., Rahman, M. A., Zhao, Z., Weiss, K., Zhang, C., Chen, Z., et al. (2018). Visualization of the cellular uptake and trafficking of DNA origami nanostructures in cancer cells. J Am Chem Soc. 140, 2478.
23. Wang, J., Li, B., Qiu, L., Qiao, X., & Yang, H. (2022). Dendrimer-based drug delivery systems: history, challenges, and latest developments. J Biol Eng. 16, 18
24. Latorre, A., Posch, C., Garcimartín, Y., Celli, A., Sanlorenzo, M., Vujic, I., et al. (2014). DNA and aptamer stabilized gold nanoparticles for targeted delivery of anticancer therapeutics. Nanoscale. 6, 7436.
25. Núñez, C., Estévez, S. V., Del Pilar Chantada, M. (2018). Inorganic nanoparticles in diagnosis and treatment of breast cancer. J Biol Inorg Chem. 23, 331.
26. Peng, H., Li, X., Zhang, H., Le, X. C. (2017). A microRNA-initiated DNAzyme motor operating in living cells. Nat Commun. 8, 14378.
27. Qiu, C., Xia, F., Zhang, J., Shi, Q., Meng, Y., Wang, C., Pang, H., Gu, L., Xu, C., Guo, Q., Wang, J. (2023). Advanced Strategies for Overcoming Endosomal/Lysosomal Barrier in Nanodrug Delivery. Research. 6, 0148

A. Appendix chapter: Extraction and concentration of viral RNA from saliva and gargle samples for the detection of SARS CoV-2

A.1 Introduction

The global COVID-19 pandemic has highlighted the vital need for fast and highly sensitive testing to effectively control, mitigate, and manage the outbreak[1, 2]. Given its highly contagious nature, even minimal exposure to just a few copies of SARS-CoV-2 may lead to infection. Therefore, the detection of even a few copies of the virus in clinical specimens is considered a positive result. Thanks to its exceptional sensitivity, reverse transcription-quantitative polymerase chain reaction (RT-qPCR) methods have been developed as the gold standard for COVID-19 testing[3]. Isothermal amplification techniques[4], particularly loop-mediated isothermal amplification (LAMP) and recombinase polymerase amplification (RPA), in combination with visual detection, simplify the process of SARS-CoV-2 detection by reducing equipment requirements and shortening the detection time and facilitate the on-site testing[5-11]. However, these techniques generally exhibit lower sensitivity compared to RT-qPCR methods. To adequately prepare for future outbreaks caused by respiratory viruses, there is a significant need to develop on-site assays with sensitivity comparable to RT-qPCR methods, which is limited by challenges for rapid and efficient extraction of viral RNA under on-site scenarios

While nasopharyngeal swabs are commonly used as clinical specimens for detecting respiratory viruses, gargle and saliva samples offer advantages over nasopharyngeal swabs due to their ease of collection, non-invasiveness, and minimal patient discomfort[7, 12]. However, the

utilization of saliva and gargle samples for virus detection does come with several technical challenges. The viscosity of saliva presents difficulties in efficiently extracting the viral RNA from samples[7, 13]. On-site viral inactivation is imperative to prevent viral transmission during transportation. The abundance of RNases in saliva can rapidly degrade the released viral RNA, necessitating measures to preserve it. Furthermore, virus loads in gargle samples are diluted compared to saliva, potentially affecting detection sensitivity[14]. Addressing these challenges requires efficient viral RNA extraction to enable on-site viral inactivation and lysis, and viral RNA preservation and concentration.

While efforts were made to develop extraction methods for viral RNA in gargle and saliva[8,9], those methods often do not meet all the requirements for on-site viral inactivation and lysis, and viral RNA preservation and concentration. Chemical lysis and heat treatments have been used for virus inactivation and lysis[8, 15], but they may not entirely mitigate the impact of RNases on viral RNA degradation[16]. Additionally, the long-time preservation of viral RNA in gargle and saliva remains understudied. Silica beads have been employed for viral RNA concentration through non-specific adsorption of nucleic acids[8, 17], which can concentrate other DNA and RNA, potentially causing interference from unrelated nucleic acids in the samples.

This work aims to develop a rapid and efficient viral RNA extraction method for on-site testing. We developed a rapid viral RNA extraction method that enables the lyses of viral particles, preserve the released viral RNA, and concentrate it in less than 10 minutes and increases sensitivity by about 20 folds. Additionally, the viral RNA can be preserved in the gargle and saliva samples with our extraction buffer for at least four weeks at room temperature. Couple with RT-qPCR, it can detect as few as 0.1 viral particle/μL in gargle samples and saliva samples.

A.2 Experimental section

A.2.1. Reagents and materials

Proteinase K and QIAamp® Viral RNA Mini Kit were purchased from QIAgen (Hilden, Germany). TaqPath™ 1-Step RT-qPCR Master Mix CG, Sodium Thiocynate, Glycogen and RNA storage solution were obtained from Thermo Scientific (Carlsbad, California, USA). CDC N1 and N2 primer-probes 2019-nCoV RUO kits were purchased from Integrated DNA Technologies™ Inc. (USA). Proteinase K Inhibitor was obtained from EMD Millipore Corp (Darmstadt, Germany). Oligo d(T)25 Magnetic Beads were purchased from New England Biolabs® Inc (Ipswich, Massachusetts, USA). TurboBeads™ Silica by Turbobeads LLC(Zürich, Switzerland), Sodium citrate, Triton-X 100, Lithium Dodecyl Sulfate (LDS), Lithium Chloride (LiCl), and Dithiothreitol (DTT) were bought from Sigma Aldrich (St. Louis, Missouri, USA). Pseudo-virus (SARS-CoV-2 RNA targets in a non-infectious viral coat) solution, AccuPlex™ SARS-CoV-2 Verification Panel-Full Genome, was obtained from Sera Care, LGC (Milford, Massachusetts, USA). DNA away, and RNase away were purchased from Thermo Fisher SCIENTIFIC. Rapid Response COVID-19 Antigen Rapid Testing Device from BTNX Inc (Onterio, Canada) is obtained at local pharmacy. Purified SARS-CoV-2 RNA was provided by our colleagues from the Department of Medical Microbiology and Immunology at the University of Alberta, the purification process was described in the previous paper.[18]

A.2.2 Preparation of spiked saliva and gargle samples.

Saliva and gargle control samples were collected from 10 healthy volunteers who had previously tested negative for SARS-CoV-2. These samples were then pooled to serve as negative gargle and saliva matrices. To prepare the spiked samples, certain amounts of pseudo-virus particles were added to aliquots of the pooled saliva or gargle samples.

A.2.3 Viral RNA extraction and concentration from saliva and gargle samples.

The process of viral RNA extraction and concentration involves two key steps: the use of the extraction buffer to release and preserve viral RNA from viral particles, followed by the concentration of the viral RNA using oligo d(T)25 magnetic beads for subsequent on-bead amplification. The optimized 1× extraction buffer comprises 0.5% w/v lithium dodecyl sulfate (LDS), 1 mM EDTA, 5 mM DTT, 0.1% w/v Tween 20, 500 mM LiCl, and 240 µg/mL proteinase K in a 100 mM Tris-HCl (pH 8.0) buffer. Specifically, 330 µL of 4× extraction buffer was first mixed with 1000 µL of saliva or gargle samples to lyse viral particles and release viral RNA. Subsequently, 150 µL of the mixture, equivalent to 113 µL of the sample, was combined with 50 µg of oligo d(T)25 magnetic beads. 150 µL of the mixture was chosen to accommodate the size of the magnetic rack, facilitating rapid separation. Additionally, this volume allows for testing samples in replicates. This solution was then incubated at room temperature for 5 minutes to capture the viral RNA onto magnetic beads through hybridization of d(T)25 with the poly (A)32 tail at the 3'-end of the viral RNA. The solution was then placed on the magnetic rack to separate the beads from the solution. The beads underwent two washing steps: first using a washing buffer (20 mM Tris-HCl, pH 7.5, 500 mM LiCl, 0.05% w/v Tween 20, 1 mM EDTA), followed by the second wash using a low-salt buffer (20 mM Tris-HCl, pH 7.5, 100 mM LiCl, 0.05% w/v Tween

20, 0.1 mM EDTA, 1× Proteinase K inhibitor). Finally, the washed beads (~5 µL) were subjected to RT-LAMP for on-bead amplification without the need for elution of the captured viral RNA.

A.2.4 Reverse transcription-quantitative polymerase chain reaction (RT-qPCR).

The one-step quantitative reverse transcription-polymerase chain reaction (RT-qPCR) is used to determine the quantity of viral RNA recovered or extracted. 15 µL of the RT-qPCR master mix comprised of 1.5 µL of N1 primer-probe mix from the U.S. Centers for Disease Control and Prevention (CDC) EUA kit (IDT cat#10006606, **Table A.1**) and 1× TaqPath™ 1-Step RT-qPCR Master Mix (Thermofisher cat# A15299, A15300), and 1 × Proteinase K inhibitor is added to the extracted RNA to a final volume of 20 µL. The reverse transcription and PCR thermocycling parameters are set per manufacturer instruction and U.S. CDC instruction of use (**Table A.2**). A standard curve was generated using a CDC-certified DNA standard from IDT (IDT cat # 10006625).

Table A.1 *N gene detection sequence and corresponding primers and probe of RT-qPCR.*[19]

Name	Sequence (5'-3')
N region	**GACCCCAAAATCAGCGAAAT**GC**ACCCCGCATTACGTTTGGTG GACC**CT**CAGATTCAACTGGCAGTAACCAGA**
Forward Primer	**GACCCCAAAATCAGCGAAAT**
Reverse Primer	**TCTGGTTACTGCCAGTTGAATCTG**
Probe	/6-FAM/**ACCCCGCATTACGTTTGGTGGACC** /BHQ1/

Table A.2 *Program of RT-qPCR.*[19]

Steps	Temperature (°C)	Duration	Cycles
UNG activation	25	2 min	N/A
Reverse Transcription	50	15 min	
Polymerase activation	95	2 min	
Denaturation	95	3 sec	50×
Extension and detection	60	30 sec	

A.2.5 Rapid antigen testing.

All samples were collected with informed consent from the volunteers, and ethics approval was obtained from the University of Alberta's Research Ethics Board. The rapid antigen test is obtained from the local pharmacy and provided for the volunteers. The collection of nasal swab or nasopharyngeal swab and rapid antigen test is done by the volunteers following the instruction provide with the kit. First morning gargle is also collected concurrently as the rapid antigen test. The results for the rapids antigen test is photographed by the volunteer and not disclosed to the researcher performing the viral RNA extraction and molecular testing on the gargle. The volunteers were asked to record the date of rapid antigen testing on the device directly using a permanent marker. Any personal identifiers were removed when possible. The volunteers were also asked to record their symptoms through out the study. The results are blinded when possible, and the results only combined at the end of the study for analysis.

A.2.6 Testing of the gargle and saliva samples.

Triplicate tests were conducted for each sample for the N gene. Each testing group included one positive control with 500 viral particles spiked into the gargle or saliva samples and one

negative control with no viral particles. After viral RNA extraction and concentration, these samples were evaluated using RT-qPCR in combination with our extraction method.

A.3 Results and discussion

A.3.1. Principle viral RNA extraction

I developed the viral RNA extraction method to accomplish three following functions: lysing the viral particles to release the viral RNA, preserving the released viral RNA, and concentrating the viral RNA onto oligo d(T)25-coated magnetic beads (Figure A.1). The development of the RNA extraction method involved two primary steps: formulating an extraction buffer to release and preserve viral RNA from viral particles and using oligo d(T)25 magnetic beads to concentrate the viral RNA for on-bead amplification. To efficiently release and preserve viral RNA, I developed an extraction buffer by combining chemical lysis and proteinase digestion. Our optimized extraction buffer consists of 0.5% w/v lithium dodecyl sulfate (LDS), 1 mM EDTA, 5 mM DTT, 0.1% w/v Tween 20, 500 nM LiCl, and 240 µg/mL proteinase K in a 100 mM Tris-HCl (pH 8.0) buffer. The LDS disrupts the viral envelope, facilitating the release of viral RNA, while proteinase K digests interfering proteins (e.g., RNases) that could degrade RNA. Additionally, DTT denatures proteins by reducing disulfide bonds, EDTA chelates Mg^{2+} required for RNase activity, Tween 20 prevents viral RNA from adhering to tube walls, and LiCl enhances subsequent concentration of viral RNA onto magnetic beads.

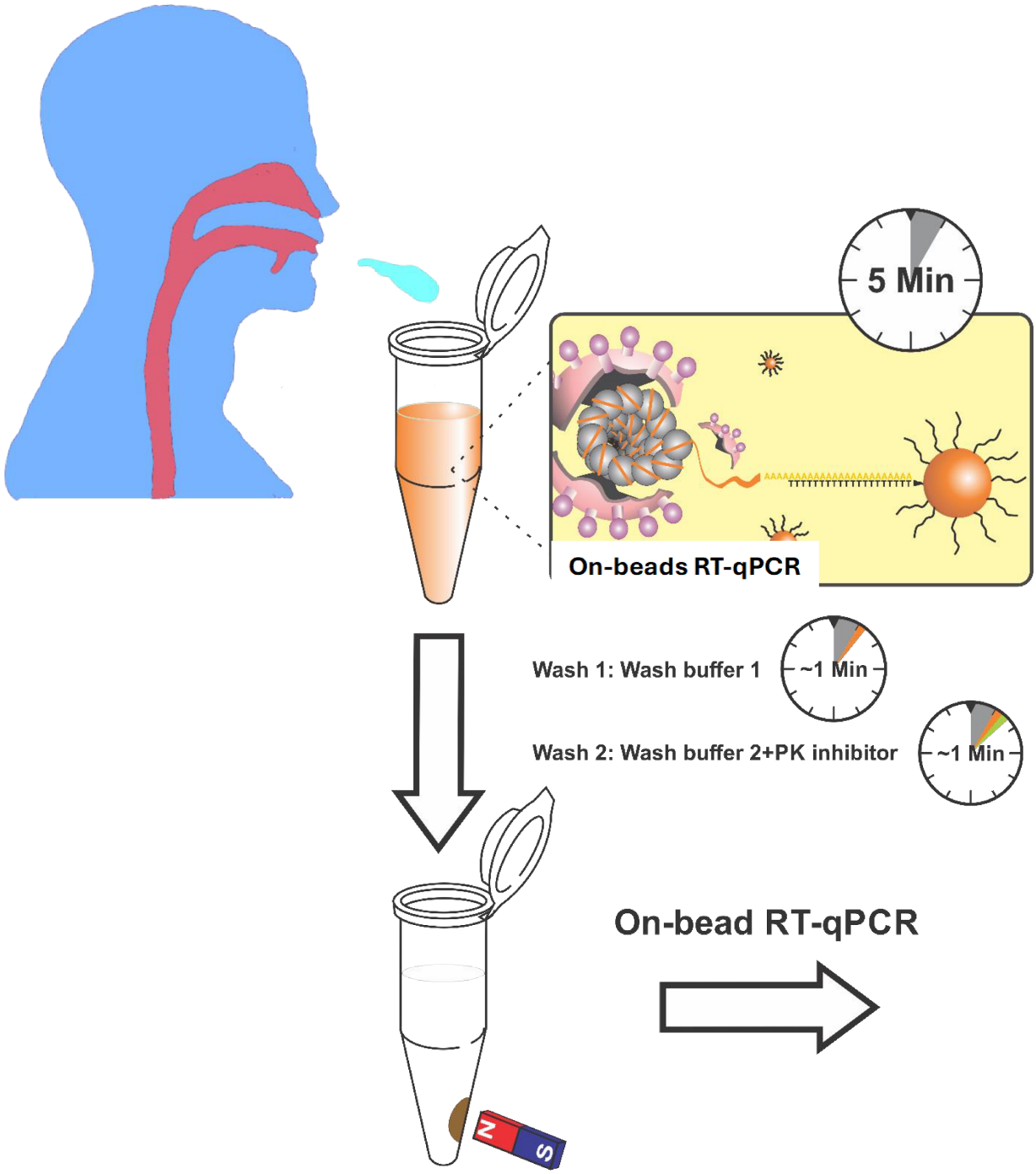


Figure A.1 *Illustration of the viral RNA extraction process using oligo d(T)25-coated magnetic bead.*

I used oligo d(T)25 magnetic beads to capture and concentrate viral RNA. These beads work by hybridization of d(T)25 on magnetic beads with the 33 adenosine residues at 3'-end of the viral RNA, which is easily accessible and free of secondary structure as the longest poly-T region is faraway and only 8nt long (nt11075-11082). Using oligo d(T)25 magnetic beads offers two advantages: Unlike other magnetic bead-based methods that capture all nucleic acid strands, this approach specifically captures viral RNA and other poly(A)-containing RNA, reducing potential interference from other nucleic acids. Additionally, the captured viral RNA can be directly used for the subsequent RT-qPCR or RT-LAMP amplification, allowing for on-bead amplification without the need for elution. This means that all captured viral RNA can be amplified, addressing a significant limitation observed in common RNA extraction kits, which typically require elution and only utilize a fraction of the eluted viral RNA for further amplification.

A.3.2. Optimization of buffer for the release of viral RNA.

I examined the performance of the viral RNA extraction method for the release, preservation, and concentration of viral RNA. We optimized the concentration of LiCl, LDS and proteinase K. I found that excess LiCl does not improve the capture efficiency of the poly-T magnetic beads extraction (**Figure A.3**), while excess LDS decreases the efficiency of the amplification (**Figure A.2**). To evaluate its efficiency in releasing and preserving viral RNA in gargle and saliva, we prepared gargle and saliva samples by adding pseudo-viral particles to pooled gargle and saliva samples collected from healthy volunteers. These samples contained 10 copies of pseudo-SARS-CoV-2 viral particles per μL. I divided each gargle and saliva sample into two groups. One group was treated with the extraction buffer with proteinase K, while the other group was subjected to the extraction buffer without proteinase K (**Figure A.4**). The sample is lysed and

incubated for 2 hours at room temperature. 100 µL of the gargle and saliva sample is extracted and concentration before quantification using RT-qPCR. Recovery is determined using a standard curve. Standard deviation shown in bracket and represented results from triplicate analyses. When analyzing the gargle samples, I observed a minor difference (≤ 0.5 cycle) in Ct value between the group treated with protease K and the group without it. However, for saliva samples, the viral RNA was significantly degraded in the group without proteinase K. These results highlight the essential role of proteinase K in preserving RNA, particularly in more complex samples such as saliva (**Figure A.4, A.5**).

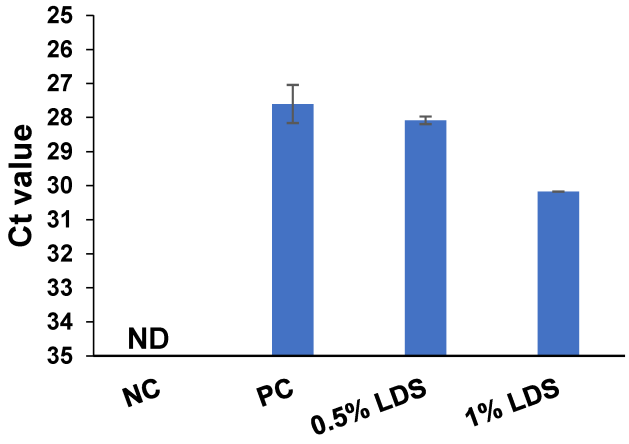


Figure A.2 *Optimization concentration of LiDS in the lysis buffer.*

Duplicate of 100 µL of saliva piked with 500 pseudo-viral particles is extracted and concentration before quantification using RT-qPCR. Negative control (NC) is non-spiked saliva. Positive control (PC) is 100 µL of purified water piked with 500 pseudo-viral particles.

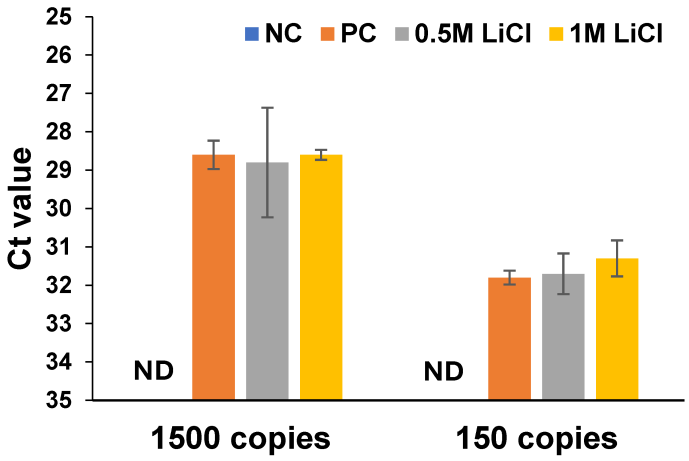


Figure A.3 *Optimization of the concentration of LiCl in the lysis buffer.*

Duplicates of 100 µL aliquots of purified viral RNA diluted in RNA storage buffer containing either 1500 or 150 copies of viral RNA is extracted and concentration before quantification using RT-qPCR. Positive controls (PC) contain purified RNA that is directly added into the RT-qPCR reaction without the extraction process.

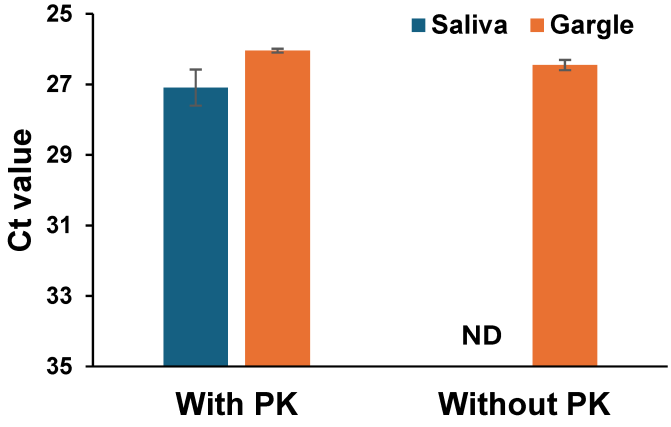


Figure A.4 *Comparison of the extraction of viral particles (VP) using lysis buffer with protease K (PK) and without protease K.*

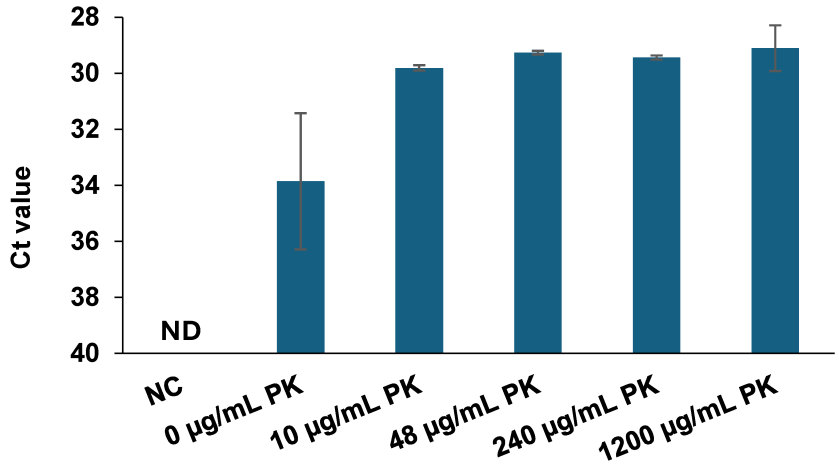


Figure A.5 *Optimization of the concentration of proteinase K (PK).*
Saliva sample is spiked with 150 viral particles 100 µL is then extracted. PK concentration does not affect the extraction efficiency significantly when >48 µg/mL

I then tested the efficiency of capturing of the released viral RNA using the oligo d(T)25 magnetic beads. We prepared two groups of gargle or saliva samples, one with viral particles at 150 copies/10µL and the other with 15 copies/10 µL. Using RT-qPCR, I quantified the extracted viral RNA and calculated recoveries. The recoveries for gargle samples were 96.0% and 96.2% for 150 and 15 copies/10 µL, respectively, and 68.9% and 80.5% for saliva samples. These high recoveries demonstrate the method can efficiently release and concentrate the viral RNA from gargle and saliva samples.

A.3.3. Preservation of viral RNA at room temperature

I also studied whether the optimized extraction buffer could preserve the released viral RNA for a long time under room temperature. I prepared two gargle or saliva samples, each containing viral particles at 50 and 5 copies/10 µL, respectively. Next, we added 4× extraction buffer to each sample at a volume ratio of 1: 3. The samples were sealed and stored at room

temperature for one month. We analyzed 150 µL aliquot of each sample at 0, 1, 2, 4, 7, 14, and 28 days. For both gargle and saliva samples, the Ct values showed no significant difference for all time points **(Figure A.6)**. Therefore, the extraction buffer can preserve the viral RNA for at least four weeks for both gargle and saliva samples. Consequently, the buffer can be preloaded into the sample collection vessels to accomplish multiple functions: viral inactivation, viral RNA release, and preservation, obviating the need for additional transport media and low temperature for delivery and storage.

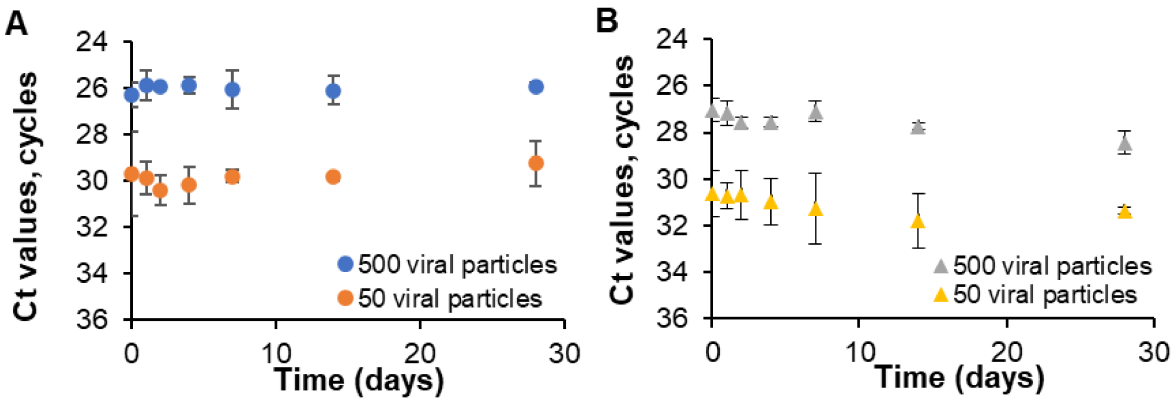


Figure A.6 *Stability of the viral RNA in saliva and gargle samples treated with the extraction buffer and stored at room temperature.*

A) gargle and B) saliva

A.3.4. Concentration of viral RNA

To examine whether our extraction method can enhance the sensitivity by concentrating the released viral RNA, we conducted a comparative analysis with the U.S. Centers for Disease Control and Prevention (CDC) method using QIAamp MinElute Virus Kits for the viral RNA extraction[19, 20]. We used the same RT-qPCR method to quantify the viral RNA extracted by two methods. For the analysis of gargle samples, our extraction method yielded a Ct value that was 4.3

cycles lower than that obtained by the U.S. CDC method, representing a sensitivity improvement of 23.6 times using a standard curve (E=105%) (**Figure A.7**). Similarly, for saliva samples, our method resulted in a Ct value 4.2 cycles smaller than that achieved by the US CDC method, corresponding to a sensitivity improvement of 20.4 times. Therefore, our extraction method can improve the sensitivity by about 20 times for both gargle and saliva samples accomplished by efficiently releasing and concentrating the viral RNA from the samples onto the magnetic beads.

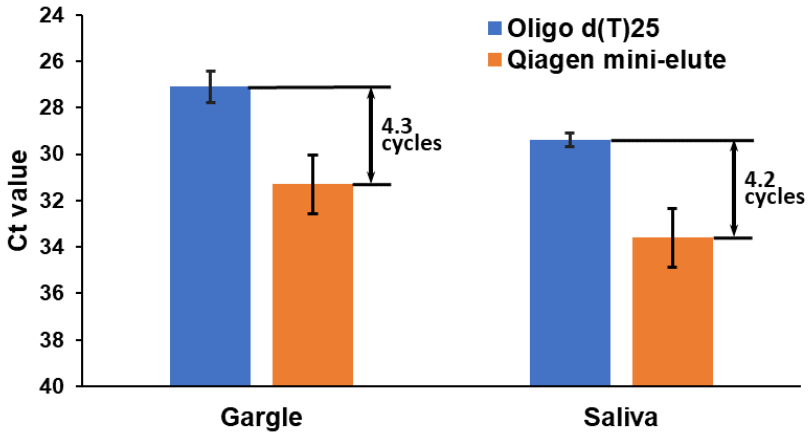


Figure A.7 *RT-qPCR results showing the capability of our viral RNA extraction method to concentrate viral RNA in comparison with the viral RNA extraction using QIAamp MinElute Virus Kits.*
100 µL of gargle or saliva samples, each containing 1500 copies of viral particles were used for this comparative analysis.

I also compared the capture efficiency and the capacity for on-beads amplification of the poly-T magnetic beads against extraction using silica-based magnetic bead (**Figure A.8 and A.9**). I extracted 200 µL of saliva spiked with 1500 copies of viral RNA using either oligo d(T) 25 MB or TurboBeads. The sample is run in duplicates. The extraction using Turbobeads is based on a previous protocol[21, 22]. I added protease K to the Turbobeads buffer to improve its extraction efficiency as Eichhoff et al. designed their original method for extraction from swab in viral transport media (Turbobeads only). After extraction and washing, we used 10 µL of 1 mM EDTA

to elute viral RNA from PolyT beads-based extraction followed by direct RT-qPCR. For the Turbobeads-extraction, viral RNA was eluted using 100 µL of RNase free water. I used 10 µL aliquot for the following RT-qPCR. The improvement in extraction between the PK (Turbo+PK) and no PK (Turbo w/o PK) group further shows the importance of PK in the treatment of biological samples. My method showed a 3.5 cycle increase compared to the improved Turbobeads-based extraction method (Turbo+PK), equal to 12-fold (E=105%) improvement in concentration. This is reasonable given that Eichhoff et al. achieved a 2-fold concentration by eluting to 100 µL. I compared the capacity of on beads by extracting viral RNA from 100µL of saliva spiked with 50 viral particles and instead of eluting the viral RNA the RT-qPCR master mix is directly mixed with the beads and vortex following washing. I can clearly see that the silica beads were unable to achieve on beads amplification.

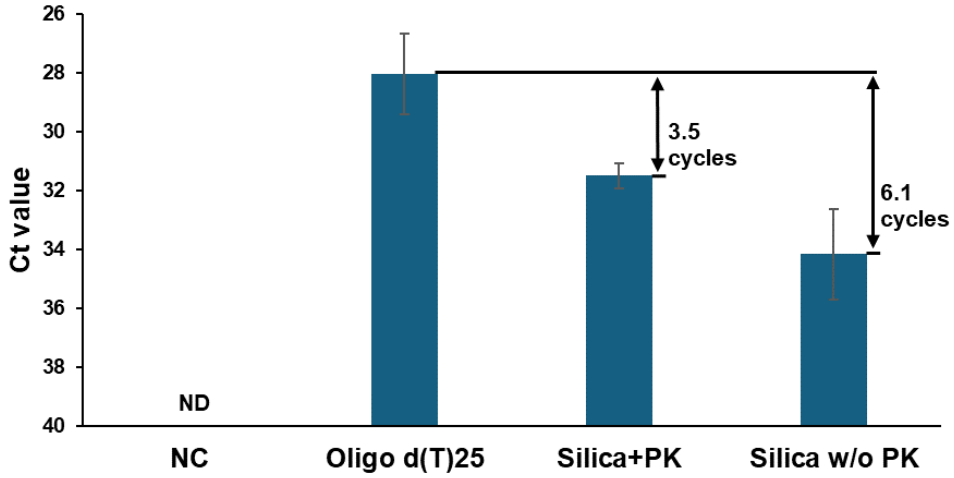


Figure A.8 *Comparison of poly-T magnetic beads (Poly T) and Turbobeads (Turbo, silica-based) for extraction of viral RNA spiked in saliva.*

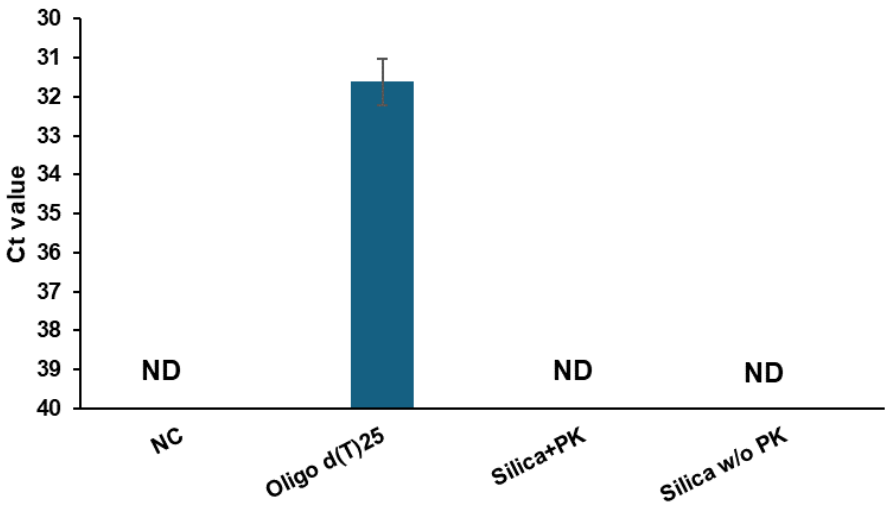
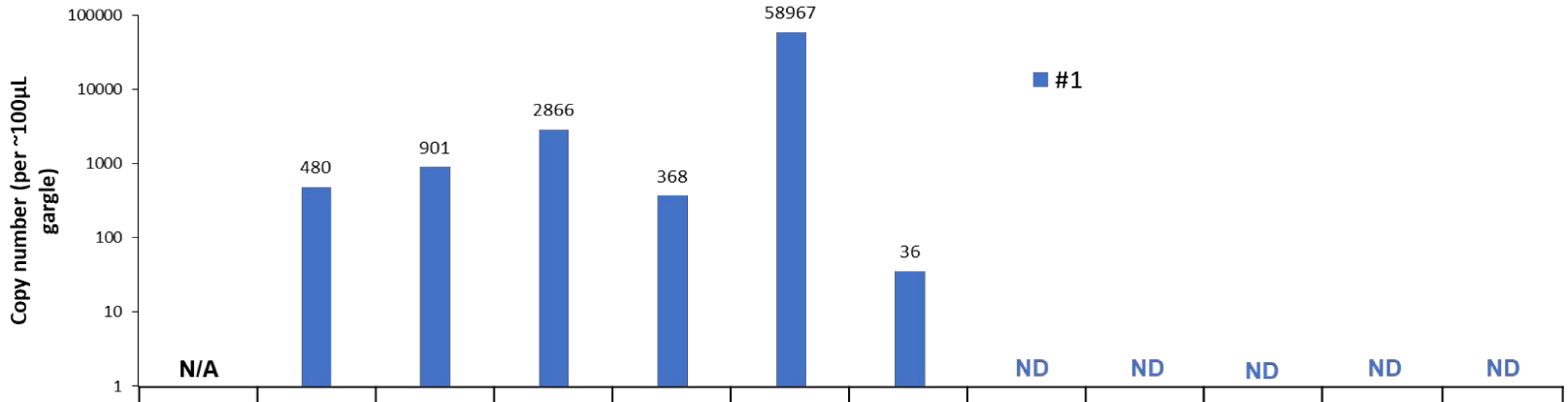


Figure A.9 *Comparison of poly-T magnetic beads(Oligo d(T)25) and Turbobeads (Silica-based) for the extraction and on-beads amplification of viral particles.*

A.3.5. Comparison of RT-qPCR coupled to our viral extraction with rapid antigen test for testing clinical samples.

To demonstrate our extraction method's applicability to clinical sample testing, I applied it to testing first-morning gargle samples from volunteers. I recruited a total of 7 volunteers and provided them with self-collection kits. I compared our viral RNA extraction coupled with RT-qPCR to rapid antigen tests. Volunteer 1 exhibited sore throat and had a negative rapid antigen test on day 0 (one day before recruitment) before testing positive on positive day 1 (time of recruitment) (**Figure A.10**). Volunteers 2 and 3, close contacts of volunteer 1 were recruited at the same time(**Figure A11, A.12**). Volunteer 2 exhibit symptoms on day 3 and viral RNA was detected RT-qPCR (**Fig A.11**). However, the rapid antigen test did not detect the infection until day 6. Volunteer 4 was recruited at a later time (**Figure A.13**). Volunteer 5 was recruited when the rapid antigen test turned positive on day 1 (**Figure A.14**). However, the volunteer reported a negative result two days prior and exhibited symptoms on day 0, suggesting that the rapid antigen test might

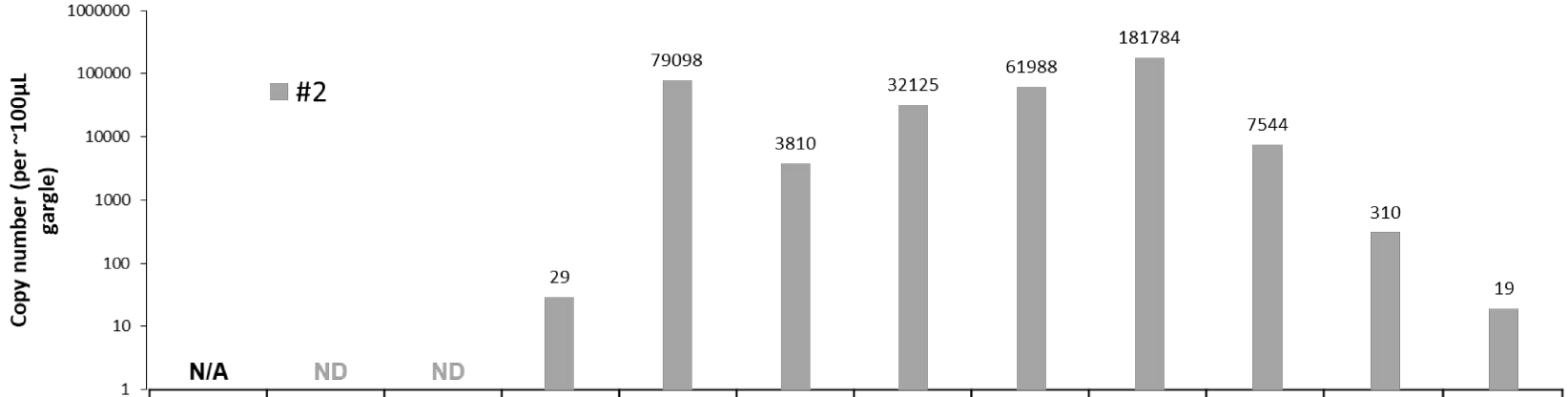
not be sensitive for early infection detection. Volunteer 6, a close contact of volunteer 5, was recruited at the same time as volunteer 5 and was asked to self-isolate. Despite our method clearly detecting viral RNA in gargle samples and the volunteer showing symptoms of sore throat and low fever, the rapid antigen test remained negative on days 1 and 2 (**Figure A.15**). Volunteer 7, who lived with volunteer 6, remained asymptomatic and tested negative for both our method and the rapid antigen test (results not shown). Results from volunteers 2 and 5 showed our method detected viral RNA before rapid antigen tests, valuable for limiting patient contact. The results from volunteers 6 and 7 underscore the importance of early detection and self-isolation following exposure and symptom onset to prevent disease spread. Volunteer 6's positive diagnosis and self-isolation effectively attenuated disease spread even when others shared the same household. Although volunteer 3's RT-qPCR result turned negative while the rapid antigen test remained positive on day 8, this finding doesn't affect patient treatment.



#1													
	Rapid antigen test												
		Neg	**Pos**	**Pos**	**Pos**	**Pos**	**Pos (faint)**	**Neg**	**Neg**	**Pos**	**Neg**	**Neg**	**Neg**
	Symptoms	Sore throat	High fever, sore throat	Low fever, sore throat	Occasional cough, loss of appetite	Occasional cough	Occasional cough	Runny nose	Occasional cough	Frequent cough	Occasional cough	Occasional cough	Occasional cough
	Total RNA (per 5 mL gargle)		**2.40E+04**	**4.50E+04**	**1.43E+05**	**1.84E+04**	**2.95E+06**	**1.78E+03**					

Figure A.10 *Results of volunteer 1.*

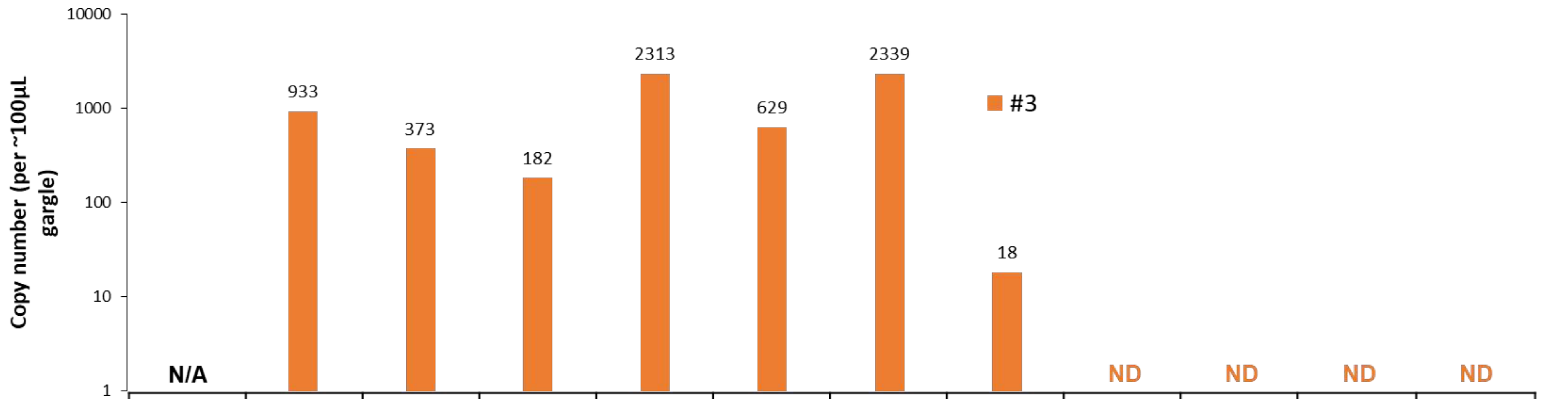
The copy number for the viral RNA obtained through poly-T MB extraction followed by RT-qPCR is shown at the top. Photograph of the rapid antigen test, the symptom and the calculated total viral load in the gargle collected is shown in the table below. Patient identifiers were removed when possible. The volunteer is asked to record the date of rapid antigen testing on the device directly using a permanent marker.



		Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11
#2	Rapid antigen test	 Neg	 Neg	 Neg	 Neg	 Neg	 Neg	 Pos	 Pos	 Pos	 Neg	N/A	 Neg
	Symptoms	Sore throat	Relieve of sore throat	Asymptom atic	Sore throat, low fever, congestion	Sore throat, low fever, lethargy, congestion	Sore throat, low fever, lethargy	Sore throat, low fever, runny nose, lethargy	Lethargy, Sore throat	Cough, Congestion	Occasional cough	Occasional cough	Occasional cough
	Total RNA (per 5 mL gargle)	N/A			1.47E+03	3.95E+06	1.91E+05	1.61E+06	3.10E+06	9.09E+06	3.77E+05	1.55E+04	9.55E+02

Figure A.11 *Results of volunteer 2.*

The copy number for the viral RNA obtained through poly-T MB extraction followed by RT-qPCR is shown at the top. Photograph of the rapid antigen test, the symptom and the calculated total viral load in the gargle collected is shown in the table below. Patient identifiers were removed when possible. The volunteer is asked to record the date of rapid antigen testing on the device directly using a permanent marker

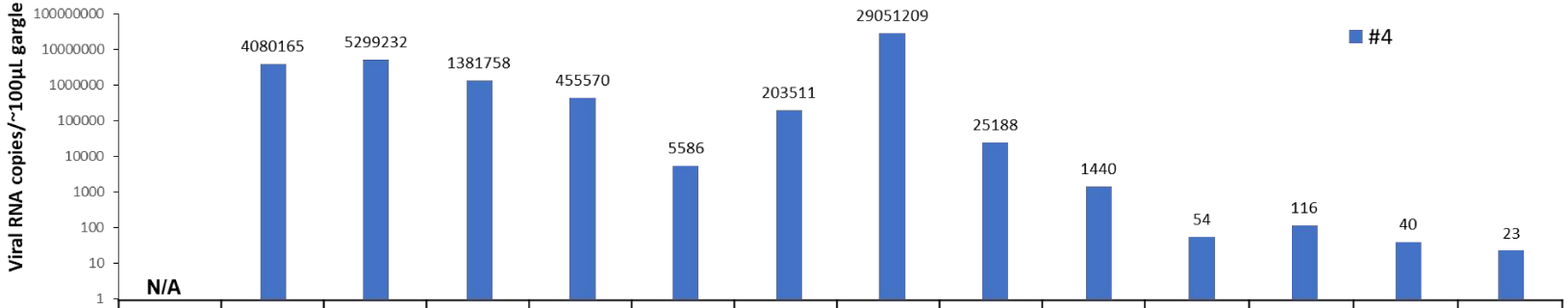


		Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11
#3	Rapid antigen test	N/A	N/A	N/A	N/A	N/A	Pos	Pos	Pos	Pos	Neg	N/A	Neg
	Symptoms	Asymptomatic	Asymptomatic	Sore throat, chills	Sore throat, low fever, lethargy, muscle pain, loss of appetite	Sore throat, low fever, lethargy, tachycardia, syncope*	Sore throat, congestion, lethargy	Cough, runny nose, congestion	Cough	Cough	Occasional cough	Occasional cough	Occasional cough
	Total RNA (per 5 mL gargle)	N/A	4.67E+04	1.87E+04	9.12E+03	1.16E+05	3.15E+04	1.17E+05	9.10E+02				

*Briefs lose consciousness after getting up in the morning

Figure A.12 *Results of volunteer 3.*

The copy number for the viral RNA obtained through poly-T MB extraction followed by RT-qPCR is shown at the top. Photograph of the rapid antigen test, the symptom and the calculated total viral load in the gargle collected is shown in the table below. Patient identifiers were removed when possible. The volunteer is asked to record the date of rapid antigen testing on the device directly using a permanent marker.



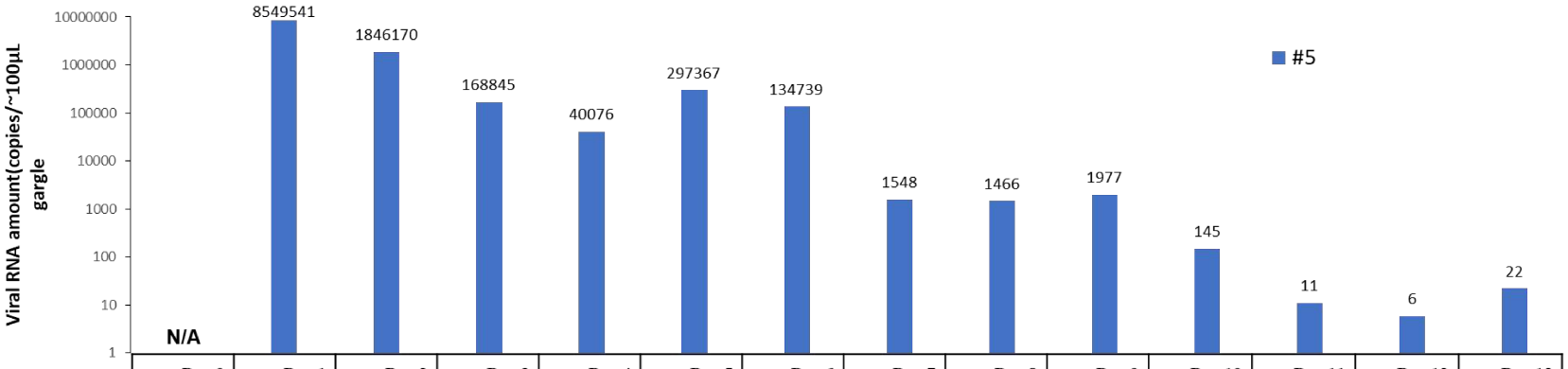
		Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13
#4	Rapid antigen test	N/A	N/A (unable to read the test strip)	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Neg	Neg	Neg
	Symptoms	Cough	Cough and lethargy	cough	Sore throat, lethargy, occasional cough	Sore throat, lethargy, occasional cough*	Sore throat, lethargy, occasional cough	throat pain, congestion, phlegm, occasional cough	throat pain, congestion, phlegm, occasional cough	throat pain, congestion, phlegm, occasional cough	Congestion, phlegmy cough, sore throat**	Congestion, cough, runny nose, slight sore throat.	minor congestion, no pain	Congestion, infrequent cough	N/A
	Total RNA (per 5 mL gargle)	N/A	2.04E+08	2.65E+08	6.91E+07	2.28E+07	2.79E+05	1.02E+07	1.45E+09	1.26E+06	7.20E+04	2.72E+03	5.79E+03	2.00E+03	1.14E+03

*Amoxicillin was taken . ** reduced redness at the throat

Figure A.13 *Results of volunteer 3.*

The copy number for the viral RNA obtained through poly-T MB extraction followed by RT-qPCR is shown at the top. Photograph of the rapid antigen test, the symptom and the calculated total viral load in the gargle collected is shown in the table below. Patient identifiers were removed when possible. The volunteer is asked to record the date of rapid antigen testing on the device directly using a permanent marker

.

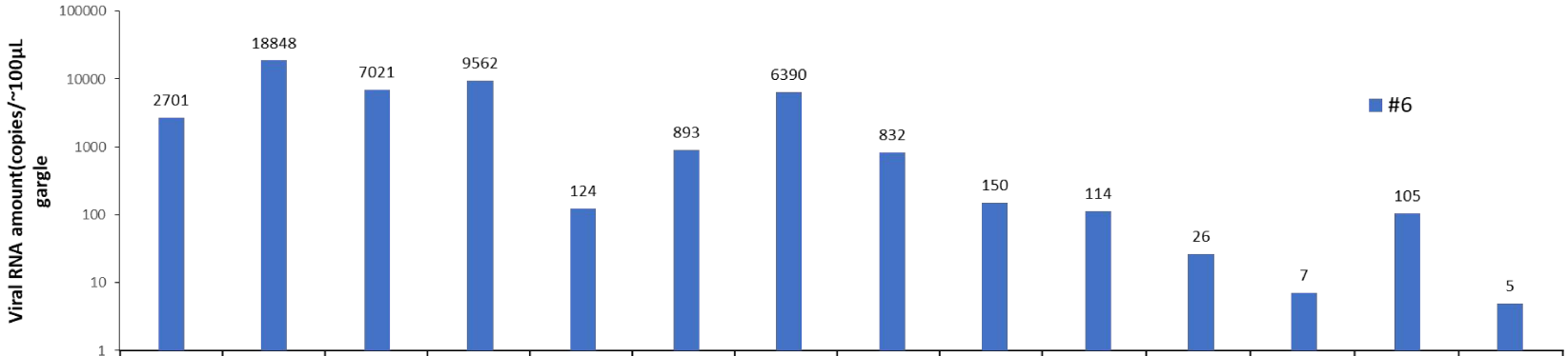


		Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13
#5	**Rapid antigen test**	**N/A** (Rapid antigen from Day -1 **Neg)**	 **Pos**	**N/A**	 **Pos**	 **Pos**	 **Pos**	 **Pos**	 **Pos**	 **Pos**(faint)	 **Pos**(faint)	 **Neg**	 **Neg**	 **Neg**	**N/A**
	Symptoms	Cough, runny nose, sore throat	Cough, runny nose, sore throat	Cough, runny nose, sore throat	Cough, runny nose, sore throat	Cough, loss of smell and taste	Complete loss of smell and taste, headache, lethargic, cough, runny nose	Cough, runny nose	Cough, runny nose	Cough, runny nose	Cough, runny nose	Occasional cough	Occasional cough	Occasional cough	N/A
	Total RNA (per 5 mL gargle)	**N/A**	**4.27E+08**	**9.23E+07**	**8.44E+06**	**2.00E+06**	**1.49E+07**	**6.74E+06**	**7.74E+04**	**7.33E+04**	**9.89E+04**	**7.25E+03**	**5.42E+02**	**2.93E+02**	**1.09E+03**

Figure A.14 *Results of volunteer 5.*

The copy number for the viral RNA obtained through poly-T MB extraction followed by RT-qPCR is shown at the top. Photograph of the rapid antigen test, the symptom and the calculated total viral load in the gargle collected is shown in the table below. Patient identifiers were removed when possible. The volunteer is asked to record the date of rapid antigen testing on the device directly using a permanent marker

.



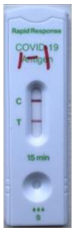
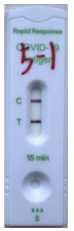
		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14
#6	Rapid antigen test														
		Neg	Neg	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Pos(faint)	Neg	Neg	Neg
	Symptoms	Swollen throat, low fever	Swollen throat	Swollen throat, dry throat, dizziness and headache	Swollen throat	Sore throat	Sore throat	Sneezing, dry cough	Sneezing, dry cough	Sneezing, cough and phlegm	Asymptomatic	Asymptomatic	Asymptomatic	Asymptomatic	Asymptomatic
	Total RNA (per 5 mL gargle)	1.35E+05	9.42E+05	3.51E+05	4.78E+05	6.21E+03	4.47E+04	3.19E+05	4.16E+04	7.48E+03	5.71E+03	1.32E+03	3.52E+02	5.27E+03	2.48E+02

Figure A.15 *Results of volunteer 6.*

The copy number for the viral RNA obtained through poly-T MB extraction followed by RT-qPCR is shown at the top. Photograph of the rapid antigen test, the symptom and the calculated total viral load in the gargle collected is shown in the table below. Patient identifiers were removed when possible. The volunteer is asked to record the date of rapid antigen testing on the device directly using a permanent marker.

A.4 Conclusion

I developed a rapid and simple method enabling extraction concentration and preservation of SARS-CoV-2 viral RNA using poly-T MB and demonstrated an application to the detection of in gargle and saliva samples. I developed an extraction buffer that enabled efficient lysis of viral particles and preservation of the released viral RNA enabling self-collection at home, which prevent crowding and need for large collection center. The released viral RNA remained stable at ambient temperature for more than 2 weeks. The extraction can be completed at room temperature in 10 minutes. I achieved near 100% recovery of viral RNA in gargle and were able to detect as few as 10 viral particles in 100 μL of saliva and gargle samples using RT-qPCR. I demonstrated an application of the method to the detection of SARS-CoV-2 RNA in gargle through comparison with the rapid antigen test testing, which demonstrated the importance of early detection of viral RNA.

A.5 References

1. Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W. C., Wang, C. B., & Bernardini, S. (2020). The COVID-19 pandemic. Crit Rev Clin Lab Sci. 57(6), 365.
2. Feng, W., Newbigging, A. M., Le, C., Pang, B., Peng, H., Cao, Y., Wu, J., Abbas, G., Song, J., Wang, D. B., Cui, M., Tao, J., Tyrrell, D. L., Zhang, X. E., Zhang, H., Le, X. C., (2020). CRISPR technology incorporating amplification strategies: molecular assays for nucleic acids, proteins, and small molecules. Anal Chem. 92(15), 10196.
3. Chu, D. K. W., Pan, Y., Cheng, S. M. S., Hui, K. P. Y., Krishnan, P., Liu, Y., Ng, D. Y. M., Wan, C. K. C., Yang, P., Wang, Q., Peiris, M., & Poon, L. L. M. (2020). Molecular Diagnosis

of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. Clin Chem. 66(4), 549.

4. Zhao, Y., Chen, F., Li, Q., Wang, L., & Fan, C. (2015). Isothermal Amplification of Nucleic Acids. Chem Rev. 115(22), 12491.
5. Baek, Y. H., Um, J., Antigua, K. J. C., Park, J.-H., Kim, Y., Oh, S., Kim, Y.-I., Choi, W.-S., Kim, S. G., Jeong, J. H., Chin, B. S., Nicolas, H. D. G., Ahn, J. Y., Shin, K. S., Choi, Y. K., Park, J. S., & Song, M. S. (2020). Development of a reverse transcription-loop-mediated isothermal amplification as a rapid early-detection method for novel SARS-CoV-2. Emerg Microbes Infect. 9(1), 998.
6. Lee, D., Chu, C. H., & Sarioglu, A. F. (2021). Point-of-Care Toolkit for Multiplex Molecular Diagnosis of SARS-CoV-2 and Influenza A and B Viruses. ACS Sens. 6(9), 3204.
7. Bellagambi, F. G., Lomonaco, T., Salvo, P., Vivaldi, F., Hangouët, M., Ghimenti, S., Biagini, D., Di Francesco, F., Fuoco, R., & Errachid, A. (2020). Saliva sampling: Methods and devices. An overview. Trends Anal Chem. 124, 115781.
8. Rabe, B. A., Cepko, C. (2020). SARS-CoV-2 detection using isothermal amplification and a rapid, inexpensive protocol for sample inactivation and purification. Proc Natl Acad Sci USA. 117(39), 24450.
9. Bokelmann, L., Nickel, O., Maricic, T., Pääbo, S., Meyer, M., Borte, S., Riesenberg, S. (2021). Point-of-care bulk testing for SARS-CoV-2 by combining hybridization capture with improved colorimetric LAMP. Nat Commun. 12(1), 1467.
10. Song, J., El-Tholoth, M., Li, Y., Graham-Wooten, J., Liang, Y., Li, J., Li, W., Weiss, S. R., Collman, R. G., Bau, H. H. (2021). Single- and Two-Stage, Closed-Tube, Point-of-Care, Molecular Detection of SARS-CoV-2. Anal Chem. 93(38), 13063.
11. Xia, S., Chen, X. (2020). Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. Cell Discov. 6(1), 37.
12. Goldfarb, D. M., Tilley, P., Al-Rawahi, G. N., Srigley, J. A., Ford, G., Pedersen, H., Pabbi, A., Hannam-Clark, S., Charles, M., Dittrick, M., Gadkar, V. J., Pernica, J. M., Hoang, L. M. N. (2021). Self-Collected Saline Gargle Samples as an Alternative to Health Care Worker-Collected Nasopharyngeal Swabs for COVID-19 Diagnosis in Outpatients. J Clin Microbiol. 59(4), e02427.

13. Landry, M. L., Criscuolo, J., Peaper, D. R. (2020). Challenges in use of saliva for detection of SARS CoV-2 RNA in symptomatic outpatients. J Clin Virol. 130, 104567.
14. Zander, J., Scholtes, S., Ottinger, M., Kremer, M., Kharazi, A., Stadler, V., Bickmann, J., Zeleny, C., Kuiper, J. W. P., Hauck, C. R. (2021). Self-Collected Gargle Lavage Allows Reliable Detection of SARS-CoV-2 in an Outpatient Setting. Microbiol Spectr. 9(1), e00361–21.
15. Davidson, J. L., Wang, J., Maruthamuthu, M. K., Dextre, A., Pascual-Garrigos, A., Mohan, S., Putikam, S. V. S., Osman, F. O. I., McChesney, D., Seville, J., Verma, M. S. (2021). A paper-based colorimetric molecular test for SARS-CoV-2 in saliva. Biosens Bioelectron. 9, 100076.
16. Green, M. R., Sambrook, J. (2019). Nested Polymerase Chain Reaction. Cold Spring Harb Protoc. doi:10.1101/pdb.prot095182
17. Juang, D. S., Juang, T. D., Dudley, D. M., Newman, C. M., Accola, M. A., Rehrauer, W. M., Friedrich, T. C., O'Connor, D. H., Beebe, D. J. (2021). Oil immersed lossless total analysis system for integrated RNA extraction and detection of SARS-CoV-2. Nat Commun. 12(1), 4317.
18. Pang, B., Xu, J., Liu, Y., Peng, H., Feng, W., Cao, Y., Wu, J., Xiao, H., Pabbaraju, K., Tipples, G., Joyce, M. A., Saffran, H. A., Tyrrell, D. L., Zhang, H., Le, X. C. (2020). Isothermal Amplification and Ambient Visualization in a Single Tube for the Detection of SARS-CoV-2 Using Loop-Mediated Amplification and CRISPR Technology. Anal Chem. 92(24), 16204.
19. Centers for Disease Control and Prevention. (2020). CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel. US Centers for Disease Control and Prevention, [cited2023 Nov. 05]. Available from: https://www.fda.gov/media/134922/download
20. Qiagen. (2020). CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel. Qiagen, [cited 2023 Nov. 05]. Available from: https://www.qiagen.com/us/products/diagnostics-and-clinical-research/sample-processing/qiaamp-viral-rna-kits/
21. Eichhoff, O. M., Bellini, E., Lienhard, R., Stark, W. J., Bechtold, P., Grass, R. N., Bosshard, P. P., & Levesque, M. P. (2020). Comparison of RNA extraction methods for the detection of SARS-CoV-2 by RT-PCR. medRxiv. https://doi.org/10.1101/2020.08.13.20172494

22. He, H., Li, R., Chen, Y., Pan, P., Tong, W., Dong, X., Chen, Y., & Yu, D. (2017). Integrated DNA and RNA extraction using magnetic beads from viral pathogens causing acute respiratory infections. Sci Rep, 7, 45199.

Bibliography

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Molecular Biology of the Cell (4th ed.). Garland Science.

Allawi, H. T., & SantaLucia, J., Jr. (1997). Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A•A, C•C, G•G, and T•T mismatches. Biochem. 36(34), 10581.

Allawi, H. T., & SantaLucia, J., Jr. (1998). Nearest-neighbor thermodynamics of internal A•C mismatches in DNA: Sequence dependence and pH effects. Biochem. 37(26), 9435.

Ang, Y. S., & Yung, L.Y. L. (2016). Rational design of hybridization chain reaction monomers for robust signal amplification. Chem Commun. 52(22), 4219.

Arora, A. A., de Silva, C. (2018). Beyond the smiley face: applications of structural DNA nanotechnology. Nano Rev Exp. 9, 1430976.

Baek, Y. H., Um, J., Antigua, K. J. C., Park, J. H., Kim, Y., Oh, S., Kim, Y. I., Choi, W. S., Kim, S. G., Jeong, J. H., Chin, B. S., Nicolas, H. D. G., Ahn, J. Y., Shin, K. S., Choi, Y. K., Park, J. S., & Song, M. S. (2020). Development of a reverse transcription-loop-mediated isothermal amplification as a rapid early-detection method for novel SARS-CoV-2. Emerg. Microbes Infect. 9(1), 998.

Bartel, D.P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell, 116, 281.

Beaucage, S. L., & Caruthers, M. H. (1981). Solid-phase synthesis of oligodeoxyribonucleotides by the phosphoramidite method. Tetrahedron Lett. 22(20), 1855.

Bellagambi, F. G., Lomonaco, T., Salvo, P., Vivaldi, F., Hangouët, M., Ghimenti, S., Biagini, D., Di Francesco, F., Fuoco, R., & Errachid, A. (2020). Saliva sampling: Methods and devices. An overview. Trends Anal Chem. 124, 115781.

Bi, S., Chen, M., Jia, X., Dong, Y., & Wang, Z. (2015). Hyperbranched hybridization chain reaction for triggered signal amplification and concatenated logic circuits. Angew Chem Int Ed. 127, 8262.

Bokelmann, L., Nickel, O., Maricic, T., Pääbo, S., Meyer, M., Borte, S., Riesenberg, S. (2021). Point-of-care bulk testing for SARS-CoV-2 by combining hybridization capture with improved colorimetric LAMP. Nat Commun. 12(1), 1467.

Broadwater, D. W., & Kim, H. D. (2012). The effect of basepair assited on DNA strand displacement. Biophys J. 102(12), 2915.

Cangialosi, A., Yoon, C., Liu, J., Huang, Q., Guo, J., Nguyen, T. D., Gracias, D. H., & Schulman, R. (2017). DNA sequence-directed shape change of photopatterned hydrogels via high-degree swelling. Science. 357, 1126.

Carr, C.E., Marky, L.A. (2017) Investigation of the melting behavior of DNA three-way junctions in the closed and open states. Biophys J. 113, 529.

Caruthers, M. H. (1985). Chemical synthesis of DNA and RNA: An overview. Bioorg Chem, 13(3), 217

Centers for Disease Control and Prevention. (2020). CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel. US Centers for Disease Control and Prevention. [cited2023 Nov. 05]. Available from: https://www.fda.gov/media/134922/download

Chen, X., Briggs, N., McLain, J. R., Ellington, A. D. (2013) Stacking nonenzymatic circuits for high signal gain. Proc Natl Acad Sci USA. 110, 5386.

Chen, Y., Gong, X., Gao, Y., Shang, Y., Shang, J., Yu, S., Li, R., He, S., Liu, X., & Wang, F. (2021). Bioorthogonal regulation of DNA circuits for smart intracellular microRNA imaging. Chem Sci. 12(47), 15710.

Chu, D. K. W., Pan, Y., Cheng, S. M. S., Hui, K. P. Y., Krishnan, P., Liu, Y., Ng, D. Y. M., Wan, C. K. C., Yang, P., Wang, Q., Peiris, M., & Poon, L. L. M. (2020). Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. Clin Chem. 66(4), 549.

Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W. C., Wang, C. B., & Bernardini, S. (2020). The COVID-19 pandemic. Crit Rev Clin Lab Sci. 57(6), 365.

Condrat, C. E., Thompson, D. C., Barbu, M. G., Bugnar, O. L., Boboc, A., Cretoiu, D., Suciu, N., Cretoiu, S. M., & Voinea, S. C. (2020). miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. Cells. 9(2), 276

Crick, F. H., Barnett, L., Brenner, S., & Watts-Tobin, R. J. (1961). General Nature of the Genetic Code for Proteins. Nature. 192(4809), 1227.

Das, G., & Lyngdoh, R. H. D. (2012). Role of wobble base pair geometry for codon degeneracy: purine-type bases at the anticodon wobble position. J Mol Model. 18(8), 3805.

Davidson, J. L., Wang, J., Maruthamuthu, M. K., Dextre, A., Pascual-Garrigos, A., Mohan, S., Putikam, S. V. S., Osman, F. O. I., McChesney, D., Seville, J., Verma, M. S. (2021). A paper-based colorimetric molecular test for SARS-CoV-2 in saliva. Biosens Bioelectron. 9, 100076.

Dirks, R. M., Pierce, N. A. (2004) Triggered amplification by hybridization chain reaction. Proc Natl Acad Sci USA. 101, 15275.

Dirks, R. M., Bois, J. S., Schaeffer, J. M., Winfree, E., & Pierce, N. A. (2007). Thermodynamic analysis of interacting nucleic acid strands1. SIAM Review. 49(1), 65.

Dong, H., Lei, J., Ding, L., Wen, Y., Ju, H., Zhang, X. (2013) MicroRNA: Function, detection, and bioanalysis. Chem Rev. 113, 6207.

Duffy K., Arangundy-Franklin, S., & Holliger, P. (2020). Modified nucleic acids: replication, evolution, and next-generation therapeutics. BMC Biology. 18, 1122.

Eichhoff, O. M., Bellini, E., Lienhard, R., Stark, W. J., Bechtold, P., Grass, R. N., Bosshard, P. P., & Levesque, M. P. (2020). Comparison of RNA extraction methods for the detection of SARS-CoV-2 by RT-PCR. medRxiv. https://doi.org/10.1101/2020.08.13.20172494

Feng, W., Newbigging, A. M., Le, C., Pang, B., Peng, H., Cao, Y., Wu, J., Abbas, G., Song, J., Wang, D. B., Cui, M., Tao, J., Tyrrell, D. L., Zhang, X. E., Zhang, H., Le, X. C.(2020). CRISPR technology incorporating amplification strategies: molecular assays for nucleic acids, proteins, and small molecules. Anal Chem. 92(15), 10196.

Feng, W., Newbigging, A. M., Tao, J., Cao, Y., Peng, H., Le, C., Zhang, H., & Le, X. C. (2021). CRISPR technology incorporating amplification strategies: molecular assays for nucleic acids, proteins, and small molecules. Chem Sci. 12(3), 1069.

Fern, J., & Schulman, R. (2018). Modular DNA strand-displacement controllers for directing material expansion. Nat Commun, 9, 3766

Fornace, M. E., Huang, J. C., Newman, T., Porubsky, N. J., Pierce, M. B., Pierce, N. A. (2022) NUPACK: analysis and design of nucleic acid structures, devices, and systems. ChemRxiv. https://api.semanticscholar.org/CorpusID:264402817

Franklin, R. E., & Gosling, R. G. (1953). Molecular Configuration in Sodium Thymonucleate. Nature. 171(4356), 740.

Gelfand, C. A., Plum, G. E., Grollman, A. P., Johnson, F., & Breslauer, K. J. (1998). Thermodynamic consequences of an abasic lesion in duplex DNA are strongly dependent on base sequence. Biochem. 37(20), 7321.

George, A. K., & Singh, H. (2016). Enzyme-free scalable DNA digital design techniques: a review. IEEE Transactions on Nanobioscience. 15(8), 928-938

Goldfarb, D. M., Tilley, P., Al-Rawahi, G. N., Srigley, J. A., Ford, G., Pedersen, H., Pabbi, A., Hannam-Clark, S., Charles, M., Dittrick, M., Gadkar, V. J., Pernica, J. M., Hoang, L. M. N. (2021). Self-Collected Saline Gargle Samples as an Alternative to Health Care Worker-Collected Nasopharyngeal Swabs for COVID-19 Diagnosis in Outpatients. J Clin Microbiol. 59(4), e02427.

Green, C., & Tibbetts, C. (1981). Reassociation rate limited displacement of DNA strands by branch migration. Nucleic Acids Res. 9(8), 1905

Green, M. R., Sambrook, J. (2019). Polymerase Chain Reaction. Cold Spring Harb. Protoc. 2019. pdb.top101857

Green, S. J., Lubrich, D., & Turberfield, A. J. (2006). DNA hairpins: fuel for autonomous DNA devices. Biophys J. 91(8), 2966.

Haley, N. E. C., Ouldridge, T. E., Mullor Ruiz, I., Geraldini, A., Louis, A. A., Bath, J., & Turberfield, A. J. (2020). Design of hidden thermodynamic driving for non-equilibrium systems via mismatch elimination during DNA strand displacement. Nat Commun. 11(1), 2461

He, H., Li, R., Chen, Y., Pan, P., Tong, W., Dong, X., Chen, Y., & Yu, D. (2017). Integrated DNA and RNA extraction using magnetic beads from viral pathogens causing acute respiratory infections. Sci Rep. 7, 45199.

Hoose, A., Vellacott, R., Storch, M., Freemont, P. S., & Ryadnov, M. G. (2023). DNA synthesis technologies to close the gene writing gap. Nat Rev Chem. 7, 144.

Jacob Souza, K., & Agrawal, D. K. (2024). Employing toehold-mediated DNA strand displacement reactions for biomedical applications. Med-X, 2(1), 13

Jafari, D., Shajari, S., & Samadikuchaksaraei, A. (2020). Modified nucleic acids: Expanding the capabilities of functional biomaterials. MDPI Molecules. 25(20), 46594

Jiang, Y.S., Li, B., Milligan, J. N., Bhadra, S., Ellington, A. D. (2013) Real-time detection of isothermal amplification reactions with thermostable catalytic hairpin assembly. J Am Chem Soc. 135, 7430.

Juang, D. S., Juang, T. D., Dudley, D. M., Newman, C. M., Accola, M. A., Rehrauer, W. M., Friedrich, T. C., O'Connor, D. H., Beebe, D. J. (2021). Oil immersed lossless total analysis system for integrated RNA extraction and detection of SARS-CoV-2. Nat Commun. 12(1), 4317.

Jung, C., & Ellington, A. D. (2014). Diagnostic Applications of Nucleic Acid Circuits. Acc Chem Res. 47(6), 1825.

Kabza, A. M., Kundu, N., Zhong, W., & Sczepanski, J. T. (2021). Integration of chemically modified nucleotides with DNA strand displacement reactions for applications in living systems. WIREs Nanomedicine and Nanobiotechnology. 14(2), e1700

Ke, Y., Castro, C., & Choi, J. H. (2018). Structural DNA nanotechnology: artificial nanostructures for biomedical research. Annu Rev Biomed Eng. 20, 375

Keller, A., & Linko, V. (2020). Challenges and Perspectives of DNA Nanostructures in Biomedicine. Angew Chem Int Ed. 59(37), 15818.

Kotani, S., & Hughes, W. L. (2017). Multi-arm junctions for dynamic DNA nanotechnology. J Am Chem Soc.139, 6363.

Kundert, K., & Fraser, J. S. (2020). DNA-binding proteins meet their mismatch. Nature. 585, 494.

Ladbury, J. E., Sturtevant, J. M., & Leontis, N. B. (1994). The Thermodynamics of Formation of a Three-Strand, DNA Three-Way Junction Complex. Biochem. 33(22), 6828.

Landry, M. L., Criscuolo, J., Peaper, D. R. (2020). Challenges in use of saliva for detection of SARS CoV-2 RNA in symptomatic outpatients. J Clin Virol. 130, 104567.

Lee, D., Chu, C. H., & Sarioglu, A. F. (2021). Point-of-Care Toolkit for Multiplex Molecular Diagnosis of SARS-CoV-2 and Influenza A and B Viruses. ACS Sens. 6(9), 3204.

Leontis, N. B., Kwok, W., Newman, J. S. (1991) Stability and structure of three-way DNA junctions containing unpaired nucleotides. Nucleic Acids Res. 19, 759.

Li, B., Ellington, A. D., & Chen, X. (2011). Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. Nucleic Acids Res. 39(15), e110.

Mao, C., Sun, W., Shen, Z., & Seeman, N. C. (1999). A nanomechanical device based on the B-Z transition of DNA. Nature. 397(6715), 144.

Mariana, L. Q., Rauhut, R., Lendeckel, W., Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. Science. 294, 853.

Mathews, C. K., & Van Holde, K. E.; Appling, D. et al., eds. (2012). Biochemistry (4th ed.). Toronto: Prentice Hall. 1181. ISBN 978-0-13-800464-4.

Morrison, L. E., & Stols, L. M. (1993). Sensitive fluorescence-based thermodynamic and kinetic measurements of DNA hybridization in solution. Biochem. 32(12), 3095.

Olson, X., Kotani, S., Yurke, B., Graugnard, E., & Hughes, W. L. (2017). Kinetics of DNA strand displacement systems with locked nucleic acids. J Phys Chem B. 121(11), 2594.

Ouldridge, T. E. Coarse-grained modeling of DNA and DNA nanotechnology. Ph.D. Thesis, University of Oxford, 2011

Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K. (2011). Structural, mechanical and thermodynamic properties of a coarse-grained DNA model. J Chem Phys. 134, No. 085101.

Ouldridge,T. E., Louis, A. A., Doye, J. P. K. (2010). DNA nanotweezers studied with a coarse-grained model of DNA. Phys Rev Lett. 104, 178101.

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., & Hofreiter, M. (2004). Genetic analyses from ancient DNA. Annu Rev Genet. 38, 645.

Pang, B., Xu, J., Liu, Y., Peng, H., Feng, W., Cao, Y., Wu, J., Xiao, H., Pabbaraju, K., Tipples, G., Joyce, M. A., Saffran, H. A., Tyrrell, D. L., Zhang, H., Le, X. C. (2020). Isothermal Amplification and Ambient Visualization in a Single Tube for the Detection of SARS-CoV‐2 Using Loop-Mediated Amplification and CRISPR Technology. Anal Chem. 92(24), 16204.

Panyutin, I. G.; Hsieh, P. (1994). The kinetics of spontaneous DNA branch migration. Proc Natl Acad Sci USA. 91, 2021.

Peng, H., Newbigging, A. M., Reid, M. S., Uppal, J. S., Xu, J., Zhang, H., & Le, X. C. (2019). Signal Amplification in Living Cells: A Review of microRNA Detection and Imaging. Anal Chem. 91(24), 15247.

Peng, H., Newbigging, A. M., Wang, Z., Tao, J., Deng, W., Le, X. C., & Zhang, H. (2018). DNAzyme-Mediated Assays for Amplified Detection of Nucleic Acids and Proteins. Anal Chem. 90(1), 190.

Peyret, N., Seneviratne, P. A., Allawi, H. T., & SantaLucia, J., Jr. (1999). Nearest-neighbor thermodynamics of deoxyribo-oligonucleotides with single base bulges. Nucleic Acids Res. 27(19), 3718

Qiagen. (2020). CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel. Qiagen, [cited 2023 Nov. 05]. Available from: https://www.qiagen.com/us/products/diagnostics-and-clinical-research/sample-processing/qiaamp-viral-rna-kits/

Qian, L., & Winfree, E. (2011). A simple DNA gate motif for synthesizing large-scale circuits. J R Soc Interface. 8, 1281.

Qian, L., & Winfree, E. (2011). Scaling up digital circuit computation with DNA strand displacement cascades. Science. 332(6034), 1196.

Qian, L., Winfree, E., & Bruck, J. (2011). Neural network computation with DNA strand displacement cascades. Nature. 475, 368.

Qiu, C., Xia, F., Zhang, J., Shi, Q., Meng, Y., Wang, C., Pang, H., Gu, L., Xu, C., Guo, Q., Wang, J. (2023) Advanced Strategies for Overcoming Endosomal/Lysosomal Barrier in Nanodrug Delivery. Research. 6, 0148.

Rabe, B. A., Cepko, C. (2020). SARS-CoV-2 detection using isothermal amplification and a rapid, inexpensive protocol for sample inactivation and purification. Proc Natl Acad Sci USA. 117(39), 24450.

Reynaldo, L. P., Vologodskii, A. V., Neri, B. P., Lyamichev, V. I. (2000). The kinetics of oligonucleotide replacements. J Mole Biol. 297(2), 511.

Ririe, K. M., Rasmussen, R. P., & Wittwer, C. T. (1997). Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal Biochem. 245(2), 154.

Rossetti, G., Dans, P. D., Gomez-Pinto, I., Ivani, I., Gonzalez, C., & Orozco, M. (2015). The structural impact of DNA mismatches. Nucleic Acids Res. 43(8), 4309

SantaLucia Jr., J. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci USA. 95, 1460.

SantaLucia Jr., J., Hicks, D. (2004) The thermodynamics of DNA structural motifs. Annu Rev Biophys. 33, 415.

Schneider, C., Rasband, W. & Eliceiri, K. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 9, 671.

Seelig, G., Soloveichik, D., Zhang, D. Y., & Winfree, E. (2006). Enzyme-Free Nucleic Acid Logic Circuits. Science. 314(5805), 1585.

Seelig, G., Yurke, B., & Winfree, E. (2006). Catalyzed relaxation of a metastable DNA fuel. J Am Chem Soc. 128(30), 12211.

Seeman, N. C., & Sleiman, H. F. (2018). DNA nanotechnology. Nat Rev Mater. 3(1). 17068.

Shang, Y., Chen, Y., Wang, Q., He, Y., He, S., Yu, S., Liu, X., & Wang, F. (2022). Construction of an endogenously activated catalytic DNA circuit for highly robust in vivo microRNA imaging. Nano Today. 45, 101553.

Shin, J.S., & Pierce, N. A. (2004). A synthetic DNA walker for molecular transport. J Am Chem Soc. 126(35), 10834.

Simmel, F. C., Yurke, B., & Singh, H. R. (2019). Principles and Applications of Nucleic Acid Strand Displacement Reactions. Chem Rev. 119(10), 6326.

Song, J., El-Tholoth, M., Li, Y., Graham-Wooten, J., Liang, Y., Li, J., Li, W., Weiss, S. R., Collman, R. G., Bau, H. H. (2021). Single- and Two-Stage, Closed-Tube, Point-of-Care, Molecular Detection of SARS-CoV-2. Anal Chem. 93(38), 13063.

Srinivas, N., Ouldridge, T. E., Sulc, P., Schaeffer, J. M., Yurke, B., Louis, A. A., Doye, J. P. K., & Winfree, E. (2013). On the biophysics and kinetics of toehold-mediated DNA strand displacement. Nucleic Acids Res. 41(22), 10641.

Turberfield, A. J., Mitchell, J. C., Yurke, B., Mills, A. P., Blakey, M. I., & Simmel, F. C. (2003). DNA fuel for free-running nanomachines. Phys Rev Lett. 90(11), 118102.

Tyagi, S., & Kramer, F. (1996). Molecular beacons: Probes that fluoresce upon hybridization. Nat Biotechnol. 14(3), 303.

Venkataraman, S., Dirks, R. M., Rothemund, P. W. K., Winfree, E., & Pierce, N. A. (2007). An autonomous polymerization motor powered by DNA hybridization. Nat Nanotechnol. 2(8), 490

Vesnaver, G., Chang, C.-N., Eisenberg, M., Grollman, A. P., & Breslauer, K. J. (1989). Influence of abasic and anucleosidic sites on the stability, conformation, and melting behavior of a DNA duplex: Correlations of thermodynamic and structural data. Biochem. 86, 3614.

Volinia ,S., Calin, G.A., Liu, C.G., Ambs, S., Cimmino, A., Petrocca, F., et al. (2006) A MicroRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci USA. 103, 2257.

Wang, J., Li, B., Qiu, L., Qiao, X., & Yang, H. (2022). Dendrimer-based drug delivery systems: history, challenges, and latest developments. J Biol Eng. 16, 18.

Watson, J. D., & Crick, F. H. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature. 171(4356), 737.

Wetmur, J. G.; Davidson, N. (1968). Kinetics of renaturation of DNA. J Mol Biol. 31, 349.

Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M., & Pierce, N. A. (2017). Constrained multistate sequence design for nucleic acid reaction pathway engineering. J Am Chem Soc. 139(4), 1788.

Xia, S., Chen, X. (2020). Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. Cell Discov. 6(1), 37.

Xuan, F., & Hsing, I.M. (2014). Triggering hairpin-free chain-branching growth of fluorescent DNA dendrimers for nonlinear hybridization chain reaction. J Am Chem Soc. 136(27), 9810.

Yakovchuk, P., Protozanova, E., & Frank-Kamenetskii, M. D. (2006). Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. Nucleic Acids Res. 34(2), 564.

Yin, P., Choi, H.M.T., Calvert, C.R., Pierce, N.A. (2008) Programming biomolecular self-assembly pathways. Nature. 451, 318.

Yin, P., Yan, H., Daniell, X. G., Turberfield, A. J., & Reif, J. H. (2004). A unidirectional DNA walker that moves autonomously along a track. Angew Chem Int Ed. 43(30), 4906.

Young, B. E., & Sczepanski, J. T. (2019). Heterochiral DNA strand-displacement based on chimeric D/L-oligonucleotides. ACS Synth Biol. 8(12), 2756.

Yurke, B., Turberfield, A. J., Mills Jr, A. P., Simmel, F. C., & Neumann, J. L. (2000). A DNA-fuelled molecular machine made of DNA. Nature. 406(6796), 605.

Yurke, B., & Mills, A. P., Jr. (2003). Using DNA to power nanostructures. Genet Program Evol M. 4, 111

Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., & Pierce, N. A. (2011). NUPACK: analysis and design of nucleic acid systems. J Comput Chem. 32(1), 170.

Zander, J., Scholtes, S., Ottinger, M., Kremer, M., Kharazi, A., Stadler, V., Bickmann, J., Zeleny, C., Kuiper, J. W. P., Hauck, C. R. (2021). Self-Collected Gargle Lavage Allows Reliable Detection of SARS-CoV-2 in an Outpatient Setting. Microbiol Spectr. 9(1), e0036121.

Zhang, D. Y., Turberfield, A. J., Yurke, B., Winfree, E. (2007) Engineering entropy-driven reactions and networks catalyzed by DNA. Science. 318, 1121.

Zhang, D.Y., Winfree, E. (2009) Control of DNA strand displacement kinetics using toehold exchange. J Am Chem Soc.131, 17303.

Zhang, D. Y. (2010). Dynamic DNA Strand Displacement Circuits. California Institute of Technology. Pasadena, California

Zhang, D. Y., & Seelig, G. (2011). Dynamic DNA nanotechnology using strand-displacement reactions. Nat Chem. 3(2), 103.

Zhang, J. X., Fang, J. Z., Duan, W., Wu, L. R., Zhang, A. W., Dalchau, N., Yordanov, B., Petersen, R., Phillips, A., Zhang, D. Y. (2017). Predicting DNA hybridization kinetics from sequence. Nat Chem. 10, 91.

Zhang, L., Chu, M., Ji, C., Tan, J., & Yuan, Q. (2022). Preparation, applications, and challenges of functional DNA nanomaterials. Nano Res. 16, 3895.

Zhao, Y., Chen, F., Li, Q., Wang, L., & Fan, C. (2015). Isothermal Amplification of Nucleic Acids. Chem Rev. 115(22), 12491.