

# Kinetic error suppression of PCR

Hiroyuki Aoyanagi<sup>a</sup>, Simone Pigolotti<sup>b</sup>, Shinji Ono<sup>a</sup>, and Shoichi Toyabe<sup>a, \*</sup>

<sup>a</sup>Department of Applied Physics, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan

<sup>b</sup>Biological Complexity Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan

\*toyabe@tohoku.ac.jp

The polymerase chain reaction (PCR) is a central technique in biotechnology. Its ability to amplify a specific target region of a DNA sequence has led to prominent applications, including virus tests, DNA sequencing, genotyping, and genome cloning. These applications rely on the specificity of the primer hybridization, and therefore require effective suppression of hybridization errors. This suppression is usually based on the energetic stability of correct hybridization. The performance of this traditional approach requires a careful design of the primer sequence and a high annealing temperature and has inherent limitations, for example in terms of reaction efficiency. Here we show that, by adding a “blocker strand” to the PCR mixture, we can sculpt a kinetic barrier that complements the traditional energetic biasing. Our method drastically suppresses the replication error by PCR without compromising the reaction efficiency. It also extends the viable range of annealing temperatures and reduces design constraint of the primer sequence. Thanks to these properties, we expect our method to significantly broaden and improve the applicability of PCR. Our approach may be extended to other biotechnology including genome editing, DNA nanotechnologies, and RNA interference.

## Introduction

PCR is used in a broad, ever-expanding range of biotechnological applications<sup>1</sup>. Fidelity of PCR is determined by the specificity of the primer hybridization. In applications, mishybridization leads to unwanted consequences, such as false positives in virus tests and sequencing errors. Given the importance and widespread nature of these applications, methods for suppressing hybridization errors are crucial.

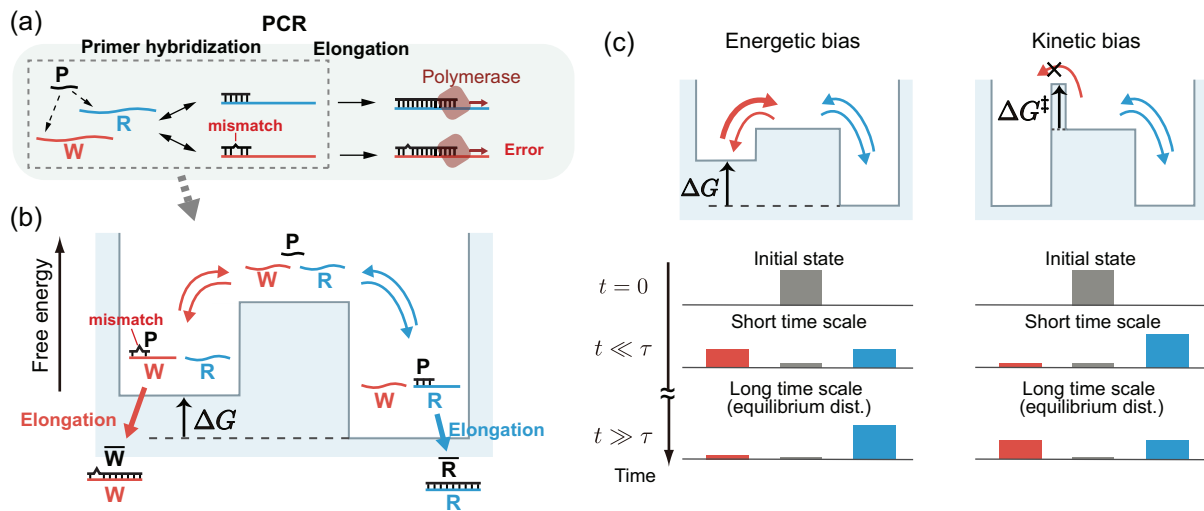
To illustrate the factors that determine the hybridization error, we consider the example of a reaction mixture containing a right template R and a contaminated wrong template W with a mutation in the primer binding region (Fig. 1a). During a PCR cycle, the temperature is lowered from a high denaturing temperature to the annealing temperature  $T_a$ . Then, a primer strand P hybridizes to either R or W. Since the primer hybridization is reversible, P repeatedly hybridizes to or dissociates from the template. Eventually, a polymerase binds to the hybridized complex P:R or P:W and elongates P to produce a complementary copy  $\bar{R}$  or  $\bar{W}$ , respectively. Important quantities characterizing the PCR are the growth rates  $\alpha_R$  and  $\alpha_W$  and error rate  $\eta$ , defined by

$$\alpha_R = \frac{r}{[R]}, \quad \alpha_W = \frac{w}{[W]}, \quad \eta = \frac{\alpha_W}{\alpha_R + \alpha_W}. \quad (1)$$

Here,  $[\cdot]$  denotes a concentration and  $r$  and  $w$  are the increases in concentrations of  $\bar{R}$  and  $\bar{W}$  in a cycle, respectively. Hence,  $\alpha_R$  and  $\alpha_W$  are the fractions of copied strands per template in a cycle. Ideally, one wants to maximize the growth rate  $\alpha_R$ , also called the PCR efficiency, and at the same time minimize the error rate  $\eta$ .

The accuracy of conventional PCR relies on primer hybridization to R, being energetically more stable than hybridization to W as quantified by the free energy difference  $\Delta G$  between P:R and P:W (Fig. 1b). This energetic bias can be increased to reduce  $\eta$  by carefully designing the primer sequence and increasing the annealing temperature  $T_a$ <sup>1</sup>.

However, this approach has an inherent limitation. To see that, we consider the hybridization kinetics (Fig. 1b). The DNA binding rate is usually diffusion-limited and thus does not significantly depend on the sequence<sup>2,3</sup>. Hence, we assume that P hybridizes to R and W at the same rate. On the other hand, the



**Figure 1. Energetic versus kinetic biasing in PCR.** (a) Standard PCR scheme. P, R, and W are the primer, right template, and wrong template with a mismatch in the primer-binding region, respectively. (b) Energy landscape corresponding to the primer hybridization in PCR. Here,  $\bar{R}$  and  $\bar{W}$  denote the right and wrong products, respectively. (c) Energy landscape for energetic (left) versus kinetic (right) biasing. At an early time,  $[P:R]$  and  $[P:W]$  are similar because of the similar barrier height for the hybridization. At equilibrium, the distribution converges to the Boltzmann distribution determined by the energetic bias  $\Delta G$ ; the amount ratio of the wrong hybridization to the right one is  $e^{-\Delta G/k_B T_a}$ , which gives  $\eta = 1/(1 + e^{\Delta G/k_B T_a})$ . Here,  $T_a$  is the annealing temperature.  $\tau$  is the relaxation time of the binding dynamics.

32 dissociation rates depend on  $\Delta G$ . Repeated hybridization and dissociation of P eventually bring the system  
 33 to thermodynamic equilibrium, where the error rate  $\eta$  is equal to  $\eta_{eq} = 1/(1 + e^{\Delta G/k_B T_a})$ . Here,  $k_B$  is the  
 34 Boltzmann constant. In the case of energetic biasing, one can show that the short time error is always larger  
 35 than  $\eta_{eq}$ <sup>4</sup>, see Fig. 1c. In fact, one problem with this approach is that the enzymatic reaction is usually quite  
 36 efficient and starts elongation before the binding equilibrates. This means that the error rate is usually not  
 37 as small as one would expect from  $\Delta G$ . Slowing down of the reaction by, for example, reducing polymerase  
 38 concentration would lower the error rate by allowing sufficient time for equilibration, but at the cost of  
 39 efficiency.

40 An alternative strategy is to sculpt a kinetic bias by building asymmetric barriers characterized by a  
 41 difference  $\Delta G^\ddagger$  so that P preferably binds to R (Fig. 1c, right). Theory predicts that such kinetic bias can  
 42 reduce  $\eta$  without sacrificing efficiency<sup>4</sup>. In this work, we make this idea concrete by introducing a “blocker”  
 43 strand in the PCR reaction mixture. We shall demonstrate that this approach improves both accuracy and  
 44 efficiency of DNA replication by PCR.

## 45 Brief methods

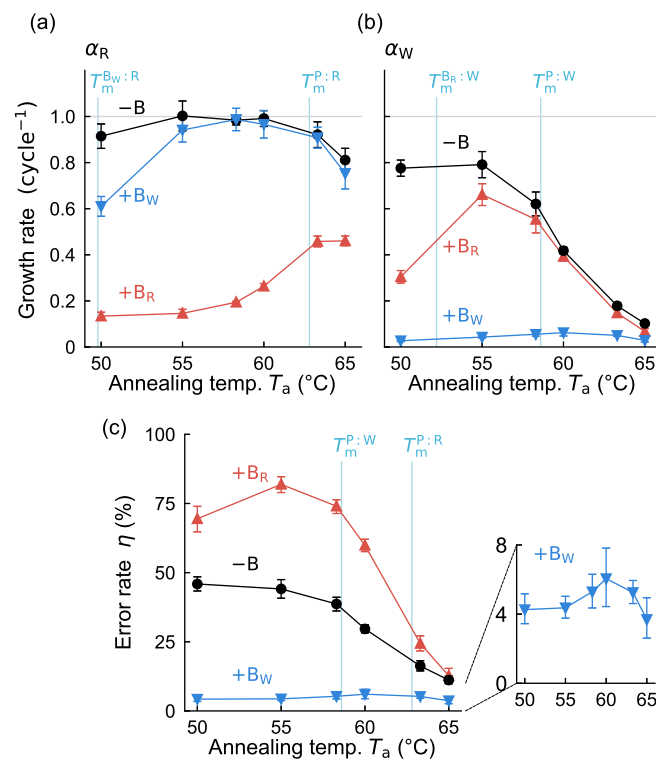
46 We perform PCR with only a single side of the primer set to focus on quantifying the error rate (Fig. S1).  
 47 Hence, the product concentration increases linearly, rather than exponentially as in the standard PCR  
 48 (Fig. S2). We mix a primer strand P, two variants of 72-nt template DNA strands R and W, indicated  
 49 concentrations of thermostable DNA polymerase, and necessary chemicals for the reactions. We also add  
 50 blocker strands depending on the experiments. The R and W templates are mixed at the same concentrations  
 51 ( $[R] = [W] = 2.5$  nM), much smaller than the primer concentration ( $[P] = 100$  nM). The P strand binds to R  
 52 without mismatches and to W with a single-base mismatch. In these conditions, the hybridization error is  
 53 expected to be large. After hybridization, polymerases copy the template and produces  $\bar{R}$  or  $\bar{W}$ . We repeat  
 54 10 or 40 thermal cycles to reduce statistical errors, measure  $r$  and  $w$ , and calculate the error rate and the  
 55 efficiency by means of Eq. (1).

56 The blocker strands  $B_R$  and  $B_W$  are 16-nt chimeric strands of DNA and locked nucleic acids (LNA)  
 57 bases. They hybridize to the primer-binding region of R and W. The  $B_{R(W)}$  strand hybridizes to R(W)  
 58 without mismatches and to W(R) with a single mismatch. Two bases at the 3' end of the blocker strands

59 are floating to prevent them from acting as primers. Blocker hybridization to the template is faster and more  
 60 stable than primer binding to the template because of their high concentration ( $[B_{R(W)}] = 20[P] = 2000\text{ nM}$ )  
 61 and the four LNA bases placed in the vicinity of the mismatch position, which significantly increases  
 62 hybridization specificity<sup>5,6</sup>.

## 63 Results

64 **PCR in the absence of blocker strands.** We first characterized the performance of conventional PCR by  
 65 measuring the efficiency  $\alpha_R$  and error rate  $\eta$  as a function of the annealing temperature. In the absence  
 66 of the blocker strands,  $\alpha_R$  and  $\alpha_W$  were large at low  $T_a$  (Fig. 2a, b). As  $T_a$  increased,  $\alpha_R$  decreased at  $T_a$   
 67 exceeding the melting temperature of P:R,  $T_m^{P:R} = 62.8^\circ\text{C}$ .  $T_m$  is defined as the temperature where half of  
 68 the DNA strands form the complex. On the other hand,  $\alpha_W$  decreased at  $T_a > T_m^{P:W} = 58.6^\circ\text{C}$ . Accordingly,  
 69 the error rate  $\eta$  was large at low  $T_a$  and decreased when  $T_a > T_m^{P:W}$  (Fig. 2c). Our results confirm that, in  
 70 conventional PCR,  $T_a$  needs to be finely tuned in the range  $T_m^{P:W} < T_a < T_m^{P:R}$  for simultaneously achieving  
 71 high  $\alpha_R$  and low  $\eta$ .

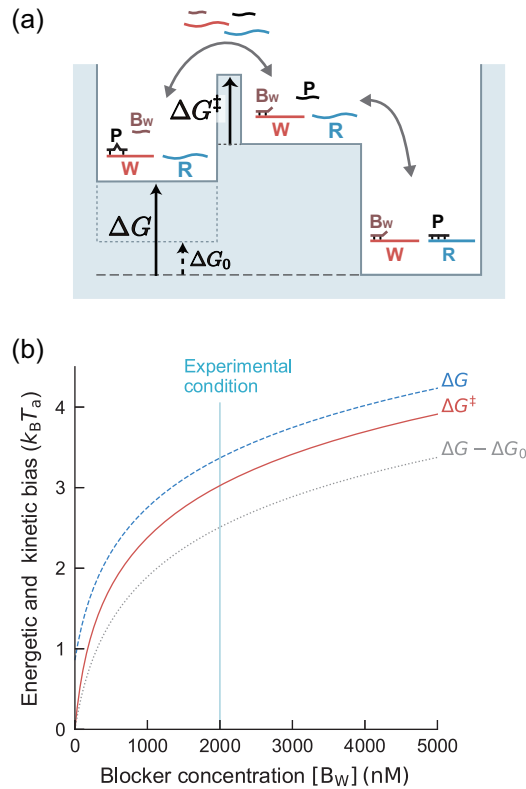


**Figure 2. PCR efficiency and error as a function of the annealing temperature.** The efficiencies of producing  $\bar{R}$  (a) and  $\bar{W}$  (b) as the function of the annealing temperature  $T_a$ . (c) Error rate  $\eta$  calculated by Eq. (1). The inset is the magnification of +B<sub>W</sub>. The error bars indicate the standard deviations. Polymerase concentration is 25 units/mL, which is the same as the standard PCR protocol.

72 **Error suppression by blocker strands.** We next studied the effect of the blocker strands on the efficiency  
 73 and the error rate. Intuitively, we expect the blockers to affect the PCR dynamics in the following way. The  
 74 blocker B<sub>W</sub> preferably hybridizes to W. As the temperature is lowered to  $T_a$  during the thermal cycle, B<sub>W</sub>  
 75 should quickly occupy most W while binding to a small fraction of R only. Hence, the hybridization of P  
 76 should be significantly biased towards R, thus suppressing the error without sacrificing the speed. On the  
 77 other hand, the addition of B<sub>R</sub> prevents P from hybridizing to R. Hence, we expect an increased error rate  
 78 in this case.

79 Indeed, the addition of  $B_W$  drastically suppressed the errors at all the annealing temperatures we tested  
 80 (Fig. 2c) without significantly impairing efficiency, at least for large  $T_a$  (Fig. 2a). At  $T_a \simeq T_m^{B_W:R} = 49.8^\circ\text{C}$ ,  
 81  $\alpha_R$  was reduced due to the hybridization of  $B_W$  to R. We found that the error suppression is still effective at  
 82  $T_a$  much lower than  $T_m^{P:W}$ , meaning that fine tuning of  $T_a$  is not needed in the presence of blockers.

83 In contrast,  $B_R$  drastically reduced  $\alpha_R$  and increased  $\eta$ . At  $T_a < 60^\circ\text{C}$ ,  $\eta$  was larger than 50 %, meaning  
 84 that  $B_R$  inverted the preference of P hybridization. This setup could be used to amplify rare sequences that  
 85 would be otherwise difficult to sample.



**Figure 3. Effective energetic and kinetic bias in the presence of blockers.**(a) The free energy landscape in the presence of  $B_W$ .  
 (b) Dependence of energetic and kinetic bias on the blocker concentration  $[B_W]$  calculated by the model (SI section S4).

86 We used chimeric DNA strands containing LNA bases for the blocker strands, which enhance the  
 87 specificity of hybridization. The blocker strands with only DNA bases had a limited effect (Fig. S6).

88 **Mathematical model.** We quantified the PCR kinetics and in particular the role of blockers using a  
 89 mathematical model (see SI section S4 for details). The model includes reversible hybridization rates of  
 90 P and  $B_R$  or  $B_W$  to the template strands. In contrast, we assume that the blockers are always at chemical  
 91 equilibrium as their high concentrations make their hybridization and dissociation dynamics very fast. For  
 92 simplicity, polymerization is modeled as a single rate without explicitly including polymerase binding and  
 93 dissociation.

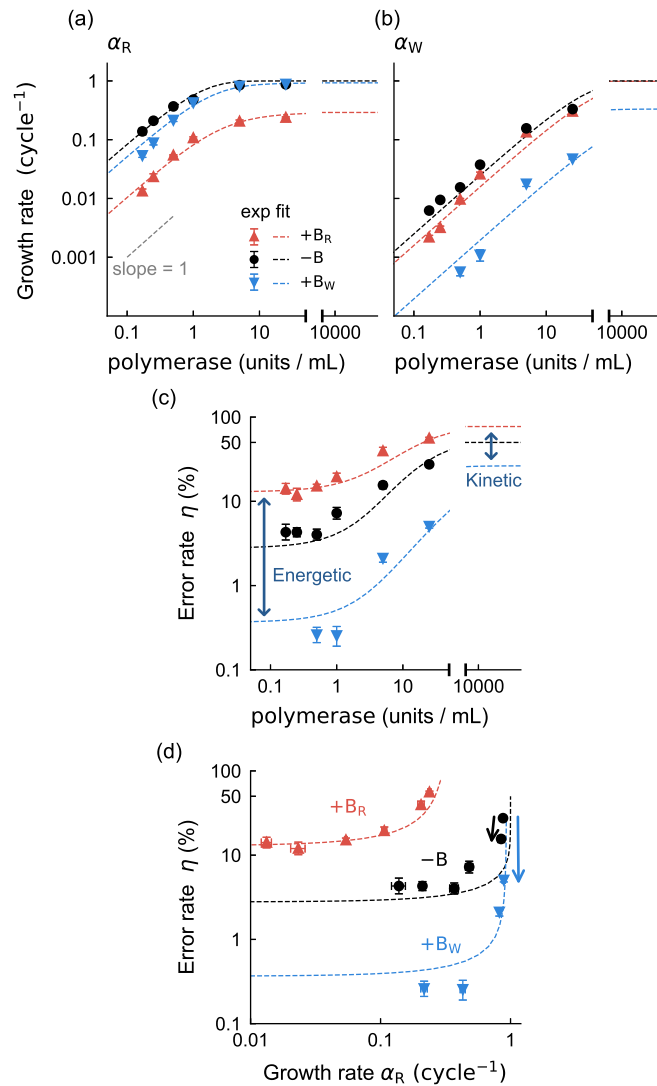
94 Introducing blockers creates an effective kinetic bias, and at the same time enhances the effective  
 95 energetic bias between right and wrong targets. These effects are quantified by

$$\Delta G^\ddagger \simeq k_B T_a \ln \left( 1 + \frac{[B_W]}{K_d^{B_W:W}} \right),$$

$$\Delta G - \Delta G_0 \simeq k_B T_a \ln \left( 1 + \frac{1}{1 + [P]/K_d^{P:W}} \frac{[B_W]}{K_d^{B_W:W}} \right). \quad (2)$$

96 Here,  $\Delta G_0$  is the energetic bias in the absence of the blocker, and  $K_d$  is the dissociation constant of the  
 97 specified hybridization. The blocker does not qualitatively alter the relation  $\Delta G > \Delta G^\ddagger$  (Fig. 3b), meaning  
 98 that the system always operates in an energetic discrimination regime<sup>4</sup>.

99 **Kinetics of error suppression.** For analyzing the detailed kinetics of error suppression by the blocker  
 100 strands, we varied the polymerase concentration by more than two orders of magnitude while fixing  $T_a$   
 101 to 60 °C, which is an appropriate temperature for our primer sequence (Fig. 4). Since elongation by  
 102 polymerase quenches the hybridization dynamics, a change in the polymerase concentration tunes the time  
 103 scale available for the hybridization dynamics to relax.



**Figure 4. Improved performance of PCR reaction in the presence of blockers is consistent with model predictions.**

Dependence of the efficiencies (a, b) and error rate (c) on the polymerase concentration. (d) The error rate is plotted against the growth rate  $\alpha_R$  (efficiency). Symbols denote the experimental data, and dashed lines correspond to the model fitting.  $\eta$  converges to the equilibrium error rate  $\eta_{eq}$  at the low polymerase concentration limit. The polymerase concentration of the standard PCR protocol is 25 units/mL.  $T_a = 60$  °C. We excluded two points with negative averages due to the statistical errors from (b) and (c) (+ $B_W$  with 0.17 and 0.25 units/mL polymerase). The error bars indicate the standard deviations. See Fig. S5 for the plots in the linear scale.

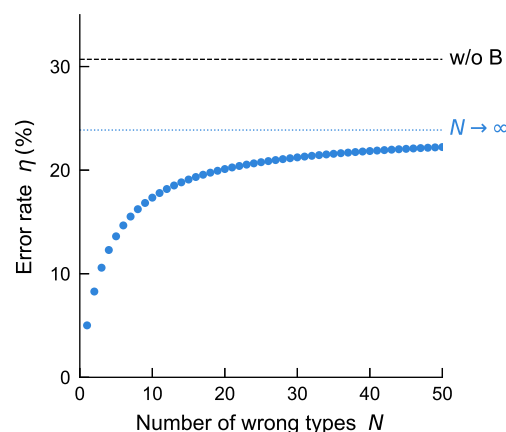
104 In the absence of blocker strands,  $\eta$  decreased as the polymerase concentration decreased (Fig. 4c).  
 105 This trends is a signature that the reaction operates in an energetic regime<sup>4</sup> as expected according to our

106 assumptions that  $\Delta G^\ddagger \simeq 0$  and  $\Delta G > 0$  (Fig. 1b). We observed similar characteristics in the presence of  $B_W$   
107 or  $B_R$  (Fig. 4c), implying that the system operates in the energetic regime even in the presence of blocker  
108 strands. This is again consistent with our prediction that the energy difference  $\Delta G$  remains larger than the  
109 energy barrier difference  $\Delta G^\ddagger$  in the presence of blockers, see Eq. 2. In the low polymerase concentration  
110 limit, the limiting value of  $\eta$  was lower in the presence of blockers. This means that the blocker addition  
111 reduces  $\eta_{eq}$  by increasing  $\Delta G$ . The error rate  $\eta$  was lower in the presence of blockers at high polymerase  
112 concentration as well, consistently with our prediction that the blockers also create an effective kinetic  
113 barrier for the wrong primer strands. These results support that the blocker strands suppress errors both  
114 energetically and kinetically as illustrated in Fig. 3a.

115 The model successfully reproduced the experimental results (dashed lines in Fig. 4). The values of the  
116 five fitting parameters were comparable with estimates based on previous work (see SI section S5).

117 Even without blocker strands, a slight reduction of polymerase concentration is effective at suppressing  
118 errors without affecting much the efficiency (black arrow in Fig. 4d). However, this strategy requires  
119 fine-tuning of the polymerase concentration to maintain the efficiency. On the other hand, the addition of  
120 blocker strands reduces errors more significantly without reducing the efficiency (blue arrow in Fig. 4d).

121 **Multiple wrong sequences.** In real-world applications of PCR, a sample may contain multiple types of  
122 unwanted sequences. We study by numerical simulations of our mathematical model if blockers could  
123 suppress replication errors in this case. For simplicity, we focus on a scenario in which the sample contains  
124  $N$  types of wrong sequences, and we add  $N$  blocker sequences, each of which perfectly hybridizes to the  
125 corresponding wrong sequence. We fix the total concentration of the wrong sequences and the blocker  
126 sequences so that the concentration of each wrong sequence and blocker sequence scales with  $1/N$ . We  
127 note that the concentration of  $B_W:W$  is roughly proportional to  $[W][B_W]$ . Since  $[W]$  and  $[B_W]$  decrease  
128 with  $N$ , blocking may become less effective with  $N$ . The error rate  $\eta$  is defined similarly to (1), but where  
129  $\alpha_W$  is the total amount of wrong products (see SI section S4). We find that, although  $\eta$  increases with  $N$ ,  
130 the blocker strands suppress  $\eta$  even in the large  $N$  limit (Fig. 5). Moreover, our model predicts that the  
131 addition of blockers should not significantly affect  $\alpha_R$  unless they strongly hybridize to R.



**Figure 5. Error rate in the presence of multiple error sequences and blockers predicted by numerical simulation.** The blue dotted line and black dashed line correspond to the result for a large number of wrong sequences ( $N = 10^4$ ) and without blocker strands, respectively.

## 132 Discussion

133 Kinetic modeling of PCR reaction has contributed to quantitatively characterize the reaction performance<sup>7-9</sup>  
134 and other aspects such as amplification heterogeneity<sup>10</sup>. However, modeling has been scarcely used to  
135 develop new guiding principles. The physics of information processing can provide such principles, thanks



136 to its progress in characterizing general biochemical reactions<sup>4,11–23</sup>. Our work demonstrates that this  
137 approach can significantly extend the performance and applicability of PCR. A similar approach can bring  
138 fruitful results when applied to other biotechnology techniques.

139 We demonstrated that adding blocker strands discriminates the right and wrong sequences by combining  
140 energetic and kinetic biasing. The kinetic biasing is effective in decreasing the error rate without affecting  
141 the efficiency. An alternative setup we studied is the use of blocker strands targeting the right template. In  
142 this case, we could increase the error rate up to a value larger than 80%. This inverted error control can not  
143 be achieved without kinetic biasing and may be helpful for sampling rare sequences.

144 Biotechnological applications of PCR are vast. Our proposed method is quite simple and therefore  
145 potentially applicable to several of these applications. The same idea may be applicable to other biotech-  
146 nology such as genome editing<sup>24</sup>, DNA nanotechnologies<sup>25</sup>, and RNA interference<sup>26</sup>, which also rely on  
147 specific hybridization of nucleic acids. In fact, a similar method that blocks unnecessary binding has been  
148 recently reported for DNA ligation<sup>27</sup>.

149 Importantly, error suppression is still effective at  $T_a$  much lower than  $T_m$  of the primer binding. This  
150 implies that we can suppress the hybridization errors in systems with limited temperature controllability,  
151 such as the hybridization inside biological cells. This might also lead to a reduced cost of applications such  
152 as virus tests by using a low-cost cycler since we do not require accurate temperature control.

153 **Acknowledgements.** ST was supported by JSPS KAKENHI Grant Numbers JP15H05460, JP18H05427,  
154 and JP19H01857. SP was supported by JSPS KAKENHI Grant Number JP18K03473 and by the Okawa  
155 Foundation (Grant Number 21-01).

## 156 **Methods**

157 **Linear PCR experiment.** DNA strands were synthesized by Eurofins Genomics, and Integrated DNA  
158 Technologies (see SI section S1). DNA/LNA chimeric strands were synthesized by Aji Bio-Pharma. The  
159 reaction mixture for polymerization contained Hot-start Taq DNA polymerase (New England Biolabs),  
160 Taq standard reaction buffer, R, W, P, and blocker strands. We performed initial heating for 30 s at 95 °C  
161 and, then, 10 or 40 cycles of 15 s at 95 °C, 30 s at 60 °C, and 5 s at 68 °C using a PCR cycler. Immediately  
162 after the cycles, the mixture was cooled down on the ice to stop the enzyme reaction and used for the  
163 quantification. The number of cycles were 40 when the polymerase concentration was 0.17, 0.25 or 0.5  
164 units/mL and 10 otherwise.

165 **Quantification of  $P_R$  and  $P_W$ .** Additional quantitative PCR was performed on a real-time PCR cycler after  
166 the linear PCR experiment for quantifying  $r$  and  $w$  (see SI section S2). The reaction mixture contains Luna  
167 Universal qPCR Master Mix (New England Biolabs), 200 nM each of the primers, and the diluted sample.  
168 The dilution rate is 1/250 in the final concentration. The thermal cycle consists of initial heating for 60 s at  
169 95 °C, 40 cycles of 15 s at 95 °C, 30 s at 66 °C, and 5 s at 72 °C.

170 **Melt curve analysis.** We measured the melting curves for the hybridization of P,  $B_R$ , and  $B_W$  to R and  
171 W from 95 °C and 20 °C and calculated their  $K_d$  (SI section S3). We mixed 100 nM each of DNA and  
172 double-strand-specific fluorescent molecule EvaGreen (Biotium). The fluorescent profile was analyzed  
173 based on the exponential background method<sup>28</sup> to obtain the melting curve.

## 174 **Data availability**

175 The data that support the findings of this study are available from the corresponding author upon reasonable  
176 request.

## 177 **Code availability**

178 The computer codes used for this study are available from the corresponding author upon reasonable  
179 request.

## 180 Author Contributions

181 HA, SP, and ST designed the research, developed the theory, and wrote the paper. HA and SO did  
182 experiments.

## 183 Competing Interests

184 The authors declare no conflict of interest.

## 185 References

- 186 1. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, 2012).
- 187 2. Tawa, K. & Knoll, W. Mismatching base-pair dependence of the kinetics of DNA-DNA hybridization studied by surface  
188 plasmon fluorescence spectroscopy. *Nucl. Acids Res.* **32**, 2372–2377 (2004). DOI 10.1093/nar/gkh572.
- 189 3. Tawa, K., Yao, D. & Knoll, W. Matching base-pair number dependence of the kinetics of DNA-DNA hybridization studied by  
190 surface plasmon fluorescence spectroscopy. *Biosens. Bioelect.* **21**, 322–329 (2005). DOI 10.1016/j.bios.2004.10.024.
- 191 4. Sartori, P. & Pigolotti, S. Kinetic versus Energetic Discrimination in Biological Copying. *Phys. Rev. Lett.* **110**, 188101 (2013).  
192 DOI 10.1103/PhysRevLett.110.188101.
- 193 5. Koshkin, A. A. *et al.* LNA (locked nucleic acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and  
194 uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* **54**, 3607–3630  
195 (1998). DOI 10.1016/s0040-4020(98)00094-5.
- 196 6. Braasch, D. A. & Corey, D. R. Locked nucleic acid (LNA): fine-tuning the recognition of DNA and RNA. *Chem. Biol.* **8**, 1–7  
197 (2001). DOI 10.1016/s1074-5521(00)00058-2.
- 198 7. Gevertz, J. L., Dunn, S. M. & Roth, C. M. Mathematical model of real-time PCR kinetics. *Biotech. Bioeng.* **92**, 346–355  
199 (2005). DOI 10.1002/bit.20617.
- 200 8. Booth, C. S. *et al.* Efficiency of the polymerase chain reaction. *Chem. Eng. Sci.* **65**, 4996–5006 (2010). DOI  
201 10.1016/j.ces.2010.05.046.
- 202 9. Bogy, G. J. & Woolf, P. J. A mechanistic model of PCR for accurate quantification of quantitative PCR data. *PLoS one* **5**,  
203 e12355 (2010). DOI 10.1371/journal.pone.0012355.
- 204 10. Best, K., Oakes, T., Heather, J. M., Shawe-Taylor, J. & Chain, B. Computational analysis of stochastic heterogeneity in PCR  
205 amplification efficiency revealed by single molecule barcoding. *Sci. Rep.* **5**, 1–13 (2015). DOI 10.1038/srep14629.
- 206 11. Peliti, L. & Pigolotti, S. *Stochastic Thermodynamics: An Introduction* (Princeton Univ. Press, 2021).
- 207 12. Schuster, P. & Eigen, M. *The Hypercycle : a principle of natural self-organization (Reprint. ed.)*. (Springer, 1979).
- 208 13. Hopfield, J. J. Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity.  
209 *Proc. Nat. Acad. Sci.* **71**, 4135–4139 (1974). DOI 10.1073/pnas.71.10.4135.
- 210 14. Ninio, J. Kinetic amplification of enzyme discrimination. *Biochimie* **57**, 587–595 (1975). DOI 10.1016/s0300-9084(75)80139-  
211 8.
- 212 15. Bennett, C. H. Dissipation-error tradeoff in proofreading. *Biosys.* **11**, 85–91 (1979). DOI 10.1016/0303-2647(79)90003-0.
- 213 16. Anderson, P. W. Suggested model for prebiotic evolution: the use of chaos. *Proc. Nat. Acad. Sci.* **80**, 3386–3390 (1983). DOI  
214 10.1073/pnas.80.11.3386.
- 215 17. Andrieux, D. & Gaspard, P. Nonequilibrium generation of information in copolymerization processes. *Proc. Nat. Acad. Sci.*  
216 **105**, 9516–9521 (2008). DOI 10.1073/pnas.0802049105.
- 217 18. Murugan, A., Huse, D. A. & Leibler, S. Discriminatory proofreading regimes in nonequilibrium systems. *Phys. Rev. X* **4**  
218 (2014). DOI 10.1103/physrevx.4.021016.
- 219 19. Sartori, P. & Pigolotti, S. Thermodynamics of error correction. *Phys. Rev. X* **5**, 041039 (2015). DOI 10.1103/phys-  
220 revx.5.041039.
- 221 20. Tkachenko, A. V. & Maslov, S. Onset of natural selection in populations of autocatalytic heteropolymers. *J. Chem. Phys.* **149**,  
222 134901 (2018). DOI 10.1063/1.5048488.
- 223 21. Toyabe, S. & Braun, D. Cooperative ligation breaks sequence symmetry and stabilizes early molecular replication. *Phys. Rev.*  
224 *X* **9**, 011056 (2019). DOI 10.1103/physrevx.9.011056.
- 225 22. Blokhuis, A., Lacoste, D. & Nghe, P. Universal motifs and the diversity of autocatalytic systems. *Proc. Nat. Acad. Sci.* **117**,  
226 25230–25236 (2020). DOI 10.1073/pnas.2013527117.
- 227 23. Rosenberger, J. H. *et al.* Self-assembly of informational polymers by templated ligation. *Phys. Rev. X* **11** (2021). DOI  
228 10.1103/physrevx.11.031055.



- 229 **24.** Anzalone, A. V., Koblan, L. W. & Liu, D. R. Genome editing with CRISPR–cas nucleases, base editors, transposases and  
230 prime editors. *Nat. Biotech.* **38**, 824–844 (2020). DOI 10.1038/s41587-020-0561-9.
- 231 **25.** Seeman, N. C. & Sleiman, H. F. DNA nanotechnology. *Nat. Rev. Mat.* **3**, 17068 (2017). DOI 10.1038/natrevmats.2017.68.
- 232 **26.** Agrawal, N. *et al.* RNA interference: Biology, mechanism, and applications. *Microbiol. Mol. Biol. Rev.* **67**, 657–685 (2003).  
233 DOI 10.1128/mubr.67.4.657-685.2003.
- 234 **27.** Gao, Y. *et al.* Accurate genotyping of fragmented DNA using a toehold assisted padlock probe. *Biosens. Bioelectron.* **179**,  
235 113079 (2021). DOI 10.1016/j.bios.2021.113079.
- 236 **28.** Palais, R. & Wittwer, C. T. Mathematical algorithms for high-resolution DNA melting analysis. *Meth. Enzym.* **454**, 323–343  
237 (2009). DOI 10.1016/s0076-6879(08)03813-5.