Avoidance of error catastrophe via proofreading innate to template-directed polymerization

Yoshiya J. Matsubara [1,*] Nobuto Takeuchi [2,3,†] and Kunihiko Kaneko [4,3,‡]

[1]*Simons Centre for the Study of Living Machines, National Centre for Biological Sciences, Bellary Road, Bangalore 560 065, Karnataka, India*
[2]*School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand*
[3]*Research Center for Complex Systems Biology, Universal Biology Institute, University of Tokyo, Komaba 3-8-1,*
Meguro-ku, Tokyo 153-8902, Japan
[4]*The Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, Copenhagen 2100-DK, Denmark*

An important issue for the origins of life is ensuring the accurate maintenance of information in replicating polymers in the face of inevitable errors. Here, we investigated how this maintenance depends on reaction kinetics by incorporating the elementary steps of polymerization into the population dynamics of polymers. We found that template-directed polymerization entails an inherent error-correction mechanism akin to kinetic proofreading, potentially generating the tolerance of long polymers to an error catastrophe at the cost of a slow polymerization process. As this mechanism does not require enzymes, it is likely to operate under broad prebiotic conditions.

I. INTRODUCTION

Template-directed polymerization is a fundamental chemical step for the sustained evolution of prebiotic systems. However, as in any chemical reaction, polymerization is subject to thermodynamically inevitable errors. Eigen [1] investigated the impact of such errors on the population dynamics of replicating polymers. Using the so-called quasispecies (QS) model, Eigen [1] showed that if the rate of error exceeds a certain threshold (i.e., *error threshold*), polymers cannot carry information in their sequence because of the presence of several incorrect (i.e., mutant) sequences, which compete with a correct (i.e., master) sequence, a phenomenon called "error catastrophe" [2–9].

Error catastrophe poses a serious issue for understanding the origin of life, because prebiotic systems most likely lack sophisticated error-correction mechanisms. By contrast, cells possess energy-driven mechanisms such as kinetic proofreading (KPR) [10–17] that increase the accuracy of template-driven replication beyond expectations due to free energy differences between correct and incorrect monomer pairs [1,10,18,19]. However, such mechanisms require multiple evolved enzymes, which may not exist in prebiotic systems. Therefore it remains unknown whether a proofreading mechanism can operating under prebiotic conditions.

————————
*yoshiyam@ncbs.res.in
†nobuto.takeuchi@auckland.ac.nz
‡kaneko@complex.c.u-tokyo.ac.jp

In this paper, we propose a prebiotic proofreading mechanism based on positive feedback between polymerization kinetics and the population dynamics of replicating templates. We propose a kinetic model of polymerization, wherein monomers are sequentially added to a primer in a template-directed manner. Using this model, we examined whether polymerization kinetics improves the tolerance of replicating templates to replication errors. We found that the effect of proofreading is maximized at the limit of infinitely fast dilution of the replicator system and, at this limit, the achievable accuracy of sequence information in replicating templates increases with the length of templates, in stark contrast with the prediction of Eigen's QS model.

II. MODEL

We consider a polymer (denoted by $X_{l,s}$) comprising a primer (denoted by "p") linked to a sequence of l binary monomers (denoted by $s \in \{0, 1\}^l$) (e.g., "p000" and "p101010") [20]. The polymer is extended by the addition of a monomer (denoted by $m \in \{0, 1\}$) using another polymer (denoted by $X_{L,S}$) as a template [Fig. 1(a)]:

$$X_{l,s} + X_{L,S} \xrightarrow{r(l,s,m,S)} X_{l+1,sm} + X_{L,S}.$$
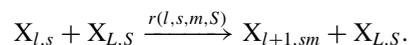
For the sake of comparison with the QS model, we assume that only polymers of length L can serve as templates (hereinafter referred to as the templates), polymers cannot be longer than L, and polymers and templates immediately separate after the addition of a monomer. The rate of monomer addition, $r(l, s, m, S)$, comprises three factors [Fig. 1(b)]:

$$r(l, s, m, S) = \beta(l, s, S)\nu(l + 1, m, S)f(S). \quad (1)$$

The first factor, $\beta(l, s, S)$, depends on the binding energy between $X_{l,s}$ and $X_{L,S}$. We assume that only a polymer bound

FIG. 1. (a) Schematic of the polymerization model (up to two joined monomers). Polymers comprise a primer (gray) and two types of monomers (blue and red). The thick arrow represents the monomer-addition reaction. In every reaction, a polymer with any sequence of length L (represented as a sequence of white monomers) is used as a template. (b) Schematic of the monomer-addition reaction using a template. In the reaction, a monomer ("0" or "1") joins a substrate polymer (e.g., "p00"), which forms pairings with a template (e.g., "p010"). The rate of the reaction is determined by the efficiency of the template f, the factor associated with the binding energy between the template and substrate β, and the joining rate of a monomer ν. β depends on whether the pairings between the substrate and the template are correct (filled circles) or not (open circles) (a pair of primers is depicted as a gray circle), and ν depends on the pairing between the added monomer and that at the corresponding position in the template.

to a template can be extended. Thus $r(l, s, m, S)$ is proportional to such fraction of $X_{l,s}$ bound to $X_{L,S}$. This fraction is expressed as $e^{-\hat{E}_a}$, where $\hat{E}_a$ is the free energy. As the binding of "correct" pairs reduces the energy, we assume $\hat{E}_a = E_0 - \hat{n}(l, s, S)\Delta$, where $\hat{n}(l, s, S)$ ($\leqslant l$) is the number of correct pairs of monomers between $X_{l,s}$ and $X_{L,S}$ [Fig. 1(b)], whereas E_0 is the energy in case of no pairings (i.e., $\hat{n} = 0$). A pair is defined as correct if the monomers are of the same type [21], and Δ ($\geqslant 0$) is the difference in free energy between correct and incorrect monomer pairs [22]. We assume the ideal condition that the fractions of the bounded polymers are small and free from saturation, that is, E_0 is large enough such that $\hat{E}_a > 0$ for any possible $\hat{n}$ ($\leqslant L$) [23]. This ideal condition [24] is also assumed in the QS model without the product-inhibition problem, as discussed later at the end of this paper. Furthermore, we rescale the time scale with e^{-E_0} for simplicity. Note that, under this condition, the time scale should be exponentially slower with the template length L. Therefore we set $\beta(l, s, S) = \exp(\hat{n}(l, s, S)\Delta)$.

The second factor in Eq. (1), $\nu(l + 1, m, S)$, depends on the binding energy between the incoming monomer m and the $(l + 1)$th monomer of the template sequence S. If this

monomer pair is correct, $\nu(l, m, S)$ is set to e^{Δ}; otherwise, it is set to unity [25]. We assume that both the monomer species, "0" and "1," exist at a constant concentration. Accordingly, $\mu = 1/(1 + e^{\Delta})$ gives the "error rate" of each monomer addition in template replication.

The third factor in Eq. (1), $f(S)$, represents the efficiency of the sequence S as a template, which can be interpreted as the fitness of S. For comparison with Eigen [1], we assume a single-peak fitness landscape: $f(S) \equiv f_0$ if $S = \{0\}^L$ (denoted by S_0 and called the master sequence); otherwise, $f(S) \equiv f_1 < f_0$ (we set $f_0 = 1$ and $f_1 = 0.1$ unless otherwise noted).

We assume a chemostat condition, where a free primer (denoted by $X_{0,p}$) was supplied at rate ϕ, and all chemical species were diluted at the same rate ϕ:

$$\varnothing \xrightarrow{\phi} X_{0,p}, \quad X_{l,s} \xrightarrow{\phi} \varnothing.$$

Thus the total concentration of all primers is unity at the steady state, $\sum_l \sum_{s \in \{0,1\}^l} x_{l,s} = 1$. We use the boundary condition in which the concentration of $X_{0,p}$ (denoted by $x_{0,p}$) is fixed, and the dilution rate ϕ is determined by $x_{0,p}$.

In summary, the equation for the rate of concentration of $X_{l,s}$ (denoted by $x_{l,s}$) is

$$\dot{x}_{l,s} = x_{l-1,s'} \sum_{S \in \{0,1\}^L} r(l-1, s', m, S)x_{L,S}$$
$$- x_{l,s} \sum_{S \in \{0,1\}^L} (r(l, s, 0, S) + r(l, s, 1, S))x_{L,S} - \phi x_{l,s},$$
(2)

where s' denotes the sequence obtained by removing the lth monomer from s and the dilution rate ϕ is determined as $\phi = \frac{x_{0,p}}{1-x_{0,p}}(1 + e^{\Delta}) \sum_{S \in \{0,1\}^L} f(S)x_{L,S}$ [26]. If $l = L$ (the maximum length), there is no second term in Eq. (2).

In the present model, the accuracy of replicated information (denoted by $\mathcal{A}$) is defined as the average similarity of all templates to the master sequence S_0 [4].

$$\mathcal{A} = \sum_{S \in \{0,1\}^L} \left(1 - 2\frac{h_{S_0,S}}{L}\right)\frac{x_{L,S}}{x_L},$$
(3)

where $x_L = \sum_{S \in \{0,1\}^L} x_{L,S}$ and $h_{S_0,S}$ is the Hamming distance between S_0 and S. For example, $\mathcal{A} = 1$ if all the templates are the master sequence, and $\mathcal{A} = 0$ if the templates are completely random sequences.

III. RESULTS

First, we demonstrate that if the dilution rate (or, equivalently, the supply of a primer) ϕ is sufficiently lower than the polymerization rate (i.e., $\phi \sim 0$ and $x_{0,p} \ll 1$), the model is reduced to the QS model. Therefore the rate equation for $x_{L,S}$ is reduced to

$$\dot{x}_{L,S} = \sum_{S' \in \{0,1\}^L} \omega_{S,S'} x_{L,S'} - \phi x_{L,S},$$
(4)

where $\omega_{S,S'} = (1 + e^{\Delta})x_{0,p} f(S')\mu^{h_{S,S'}}(1 - \mu)^{L - h_{S,S'}}$ and μ is the error rate (i.e., $\mu = \frac{1}{1+e^{\Delta}}$) (see Appendix A for the derivation). This is because, for low ϕ, a primer is extended to the

FIG. 2. (a) Dependence of the accuracy of the replicated information $\mathcal{A}$ [defined as Eq. (3)] on the fixed concentration of free primers, $x_{0,p}$ for template length $L = 6$. Different error rates μ are shown in different colors. (b) Dependence of $\mathcal{A}$ on μ for different $x_{0,p}$, as plotted with different colored curves. The black dashed curve represents the result of the QS model, whereas the solid black curve represents $\mathcal{A}$ in the case of an infinite dilution rate, $x_{0,p} \sim 1$, calculated by Eq. (8).

maximum length L immediately after binding to a template, and fast polymerization does not affect the slow population dynamics of the templates.

Hence, in this low-dilution-rate limit, our model produces the same error catastrophe and the threshold as provided by the QS model: The accuracy of the replicated information $\mathcal{A}$, defined as Eq. (3), decreases monotonically with the increase in μ, and the decrease is accelerated as the sequence length of templates L increases. Based on the correspondence of Eqs. (4) and the QS model, the error rate μ has to be smaller than the error threshold in order to maintain $\mathcal{A} \sim 1$:

$$\mu \lesssim \ln(W)/L, \qquad (5)$$

where $W \equiv \frac{f_0}{f_1}$ is the advantage of the master sequence. This is the error threshold derived by Eigen [1].

Then, we consider the situation wherein the dilution rate ϕ is comparable to the polymerization rate. In this case, the model is not reduced to the QS model, and the polymerization kinetics significantly affect the accuracy of replicated information $\mathcal{A}$. We computed $\mathcal{A}$ in the steady state in Eq. (2), varying the concentration of the free primer ($x_{0,p}$) by tuning the dilution rate ϕ [Fig. 2(a)] [27]. If $x_{0,p}$ (i.e., ϕ) is low, the accuracy $\mathcal{A}$ approaches that of the QS model, as expected [Fig. 2(b)]. As $x_{0,p}$ (i.e., ϕ) increases, $\mathcal{A}$ monotonically increases for any error rate μ. In other words, slow polymerization relative to dilution improves the accuracy $\mathcal{A}$.

The increased accuracy is attributed to the polymerization process, which works toward achieving multistep error correction for each monomer site in the template sequence. Here, we derive the maximally achievable accuracy in the limit of an infinite dilution rate ϕ (i.e., $x_{0,p} \sim 1$).

First, we consider how the concentrations of polymers depend on the monomer at a specific site. Let the concentration of the polymers of length l whose ith monomer is $m \in \{0, 1\}$ be expressed as $\xi_{l,m}^{(i)} x_l$, where $\xi_{l,0}^{(i)} + \xi_{l,1}^{(i)} = 1$, $x_l = \sum_{s \in \{0,1\}^l} x_{l,s}$, and $i \leqslant l \leqslant L$. Using Eq. (2), we can show that as $\phi \to \infty$,

$$\frac{\xi_{l,0}^{(i)}}{\xi_{l,1}^{(i)}} = \frac{F_0^{(i)}}{F_1^{(i)}} \frac{\xi_{l-1,0}^{(i)}}{\xi_{l-1,1}^{(i)}}, \qquad (6)$$

where $F_m^{(i)}$ is the relative rate at which the polymers of length $l - 1$ are extended to the polymers of length l whose ith monomer is m (see Fig. 5 in Appendix B). The values of $F_m^{(i)}$ are estimated as

$$F_0^{(i)} = e^\Delta \xi_0 f_0 + e^\Delta \big(\xi_0^{(i)} - \xi_0\big) f_1 + \xi_1^{(i)} f_1,$$
$$F_1^{(i)} = \xi_0 f_0 + \big(\xi_0^{(i)} - \xi_0\big) f_1 + e^\Delta \xi_1^{(i)} f_1, \qquad (7)$$

where $\xi_m^{(i)} = \xi_{L,m}^{(i)}$ and ξ_0 is the fraction of the master sequence, i.e., $\xi_0 = \prod_j \xi_0^{(j)}$ (see the derivation in Appendix B). To derive Eq. (7), we assumed that monomer additions at different positions in a sequence are independent of each other [28]. In both equations in (7), the first term represents the rate of monomer addition using the master sequence (i.e., the sequence with all 0s) as a template, and the second and third terms represent rates using other templates with the ith monomer "0" and "1," respectively.

Next, polymer sequences of length $i - 1$ must undergo $L + 1 - i$ steps of monomer-addition reactions in order to complete the synthesis of a template with length L. Using Eq. (6) recursively, the fractions $\xi_0^{(i)}$ and $\xi_1^{(i)}$ are derived by self-consistently solving

$$\frac{\xi_0^{(i)}}{\xi_1^{(i)}} = \frac{\big(F_0^{(i)}\big)^{L+1-i}}{\big(F_1^{(i)}\big)^{L+1-i}}. \qquad (8)$$

As shown in Fig. 2(b), we calculated the accuracy $\mathcal{A}$ using this estimate for $\xi_0^{(i)}$ as $\mathcal{A} = \frac{2}{L} \sum_i \xi_0^{(i)} - 1$, which agrees well with the simulation result for $x_{0,p} = 0.9$.

The effective error rate $\xi_1^{(i)}$ given by the solution of Eq. (8) is less than the original error rate μ as described below, which suggests a proofreading effect working. By assuming the dominance of the master sequence ($\xi_0^{(i)} \sim \xi_0$, $\xi_1^{(i)} \sim 0$), $\xi_1^{(i)}$ is approximated as

$$\xi_1^{(i)} \sim \frac{1}{1 + e^{(L+1-i)\Delta}} \sim e^{-(L+1-i)\Delta}, \qquad (9)$$

because $F_0^{(i)}/F_1^{(i)}$ is approximated by e^Δ. Equation (9) agrees with the minimum error rate that can be achieved in the KPR model with $L + 1 - i$ steps, when the binding energy between the enzyme and correct or incorrect substrate differs by Δ [10,13]. Furthermore, even if the fraction of the master sequence ξ_0 is close to zero (i.e., Δ is small), the small difference between $F_0^{(i)}$ and $F_1^{(i)}$ is amplified with the powers of $L + 1 - i$, possibly resulting in a significant difference between $\xi_0^{(i)}$ and $\xi_1^{(i)}$, and high $\mathcal{A}$ accordingly.

Given the effective error rate in Eq. (8) at each monomer site, we calculated the error threshold for the correct information and found it to be dominant. In Fig. 3(a), we have plotted the dependence of the information accuracy $\mathcal{A}$ on the error rate μ with various template lengths L. At the limit of an infinite dilution rate, the error threshold for μ, at which the information is lost ($\mathcal{A} \sim 0$), increases with the length of the template L [Fig. 3(b)] [29]. This is in sharp contrast with the QS model, wherein the fraction $\mathcal{A}$ declines sharply with L under any condition, and the error threshold for μ approaches zero with an increase in L, as expressed by Eq. (5).

The increase in accuracy of the information with length L in Fig. 3 is achieved because increasing L increases the
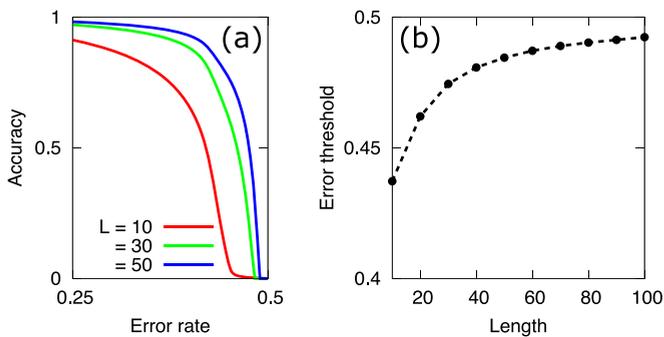
FIG. 3. (a) Maximum achievable accuracy of the information plotted as a function of the error rate μ ($\mu = \frac{1}{1+e^\Delta}$) for different lengths of polymer L in the case with $x_{0,\mathrm{p}} \sim 1$. The accuracy was calculated using Eq. (8). (b) Dependence of the error threshold on the length of template L. The threshold is calculated as the error rate μ, satisfying $\mathcal{A} = 0.25$.

number of reaction steps a monomer site in the sequence undergoes before template synthesis is completed. The effective error rate at each monomer site in the template is exponentially reduced with the number of steps, as in multistep KPR. Although the variety of incorrect sequences increases exponentially with L, as in the QS model, this can be overcome by the proofreading effect (see Appendix C).

Finally, we discuss the trade-off relationship between the accuracy and template yield. This trade-off is inevitable, because the accuracy of the KPR is generally achieved at the expense of synthesis efficiency [12,13]. In our model, we computed the yield as the actual concentration of the master sequence x_{L,S_0}. In Fig. 4(a), the yield is plotted against the accuracy of information $\mathcal{A}$ by varying the dilution rate ϕ. With an increase in ϕ, the accuracy increases, but the yield
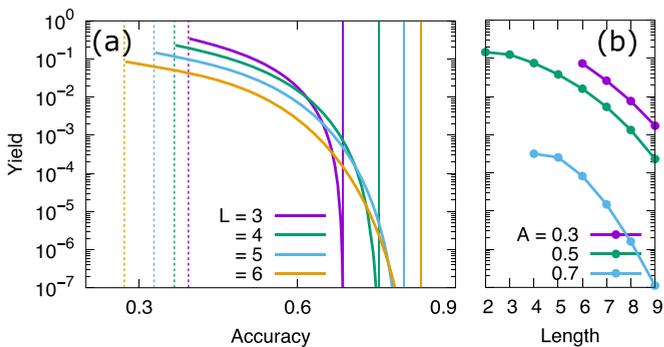
FIG. 4. (a) Trade-off between the accuracy of the replicated information $\mathcal{A}$ and the template yield defined as the master sequence concentration x_{L,S_0}. The solid curves represent the dependence of x_{L,S_0} on $\mathcal{A}$. The dilution rate ϕ is varied by controlling the fixed concentration of the free primer $x_{0,\mathrm{p}}$ in the chemostat, from ~ 0 to ~ 1. The dashed vertical lines represent the accuracy $\mathcal{A}$ in the QS model, and the solid vertical lines represent the case with $x_{0,\mathrm{p}} \sim 1$. We set $\Delta = 1$ [i.e., $\mu = 1/(1 + e)$]. (b) The yield x_{L,S_0} varies with the length of the template L at fixed accuracy $\mathcal{A} = 0.3, 0.5, 0.7$ (note that some values of $\mathcal{A}$ cannot be achieved for lower L, and so some data points are missing). The yield decreases with length L faster than exponential decay.

decreases. A similar trade-off exists between the accuracy and energy influx (see Appendix E).

IV. SUMMARY AND DISCUSSION

In summary, in our template-polymerization system, the proofreading effect reduces the effective error rate as long as the dilution is not too slow. The effective error rate decreases with the template length (in sharp contrast with the QS model) in the slow dilution limit, whereas the error increases with the length. This proofreading effect entails a trade-off between the accuracy of replicated information and the production yield, as in the KPR scheme.

We made the following four major assumptions to make our model comparable with the QS model. However, most of these assumptions can be relaxed for proofreading to work in our model.

First, we assumed immediate separation of the templates after complete full-length polymerization, as is also assumed in the QS model; thus the separation of shorter sequences is immediate based on thermodynamic reasoning [30]. Without this assumption, the so-called "product inhibition" problem occurs, as already pointed out for the QS model [31,32]. However, the product-inhibition problem could be circumvented by assuming the binding energy of monomer pairing to be weak while maintaining accurate replication due to the proofreading effect discussed in this paper; there are physicochemical conditions that are free from this problem [33]. Experimentally, this condition can be realized by the mode of driving template separation (e.g., fast environmental, thermal or tidal, or oscillation [34–37]).

Second, only the longest polymers were assumed to work as templates. However, this assumption is not essential in contrast to the QS model, if the shorter templates rebind to the longer templates and are consumed to produce them [38]. Even if the shorter polymers act as templates, our results do not significantly differ over certain parameter regions (see Appendix F). A similar effect has also been observed in experimental templated ligation systems [37]. Because the proofreading in our model works better for a larger L, our results suggest that a mechanism for selecting longer polymers (e.g., Refs. [39,40]) would resolve the error-catastrophe problem because of the proofreading effect.

Third, we considered the simplest "fitness landscape" $f(S)$, wherein only the master sequence has high fitness. Here, "fitness" represents the efficiency of a template for incorporating monomers. It is also possible to consider the arbitrary fitness landscape in our model (e.g., a multimodal or more rugged landscape) [4,6]. We examined a few alternate landscapes, which supported that the proofreading effect is relevant to avoid the error catastrophe (see Appendix G) [41].

Fourth, we assumed that sequences always bind to templates at the same position. However, our proofreading mechanism holds even if sequences could rebind to any site in a template; that is, rebinding to wrong locations is insignificant because such cases are rare due to a small number of correct pairings. In most cases, a polymer extends by binding to the correct site resulting in more correct pairings.

However, our proofreading scheme has a few limitations in cases with longer template lengths. Although the accuracy

$\mathcal{A}$ monotonically increases with length L under the conditions assumed in our study [Fig. 3(b)], the time scale to complete the replication reaction increases and the yield of the product drastically decreases with L [Fig. 4(b)]. We assumed the ideal condition wherein the bond between polymers separates immediately after the reactions, and no saturation of the fraction of bounded polymers exists. To achieve this condition, the time scale of the reactions must be exponentially slower with an increase in the template length, because the unbinding of longer polymers would take a longer time. This corresponds to the rescaling of the time by e^{-E_0} as explained in the model description, where E_0 must be larger than ΔL. In addition, as we assumed fast dilution during polymerization, the product template yield decreases faster than exponential with its length, as highlighted in Fig. 4(b). Therefore, with a sufficiently long template, the reaction speeds and yields of products could be too low to work in real experimental systems.

Note, however, that such a limitation could be avoided by considering a more realistic and detailed model. As discussed above, sites in a template closer to the primer undergo more error-correction steps, resulting in fewer errors in the product template. The proofreading effects for those sites are excessive. In this paper, we consider the ideal condition where a mismatch at every site equally affects the extension rate of the sequence, which entails a long time scale for the reactions. However, if a more realistic and detailed model is considered wherein the effect of mismatches at each site on the extension rate depends on their proximity to the ends of the sequence, the reduction of errors while increasing the yields of production and speeds of the replication process up to the realistic regime might be maintained.

In principle, our scheme works even in synthetic replicating systems without complex reaction pathways such as a nonenzymatic primer-extension system [42,43] or a template-directed ligation system [37] (see Refs. [44,45] for reviews). Finally, we briefly compare other schemes with the proposed model. The standard KPR currently used in biological systems requires the specific design of the reactions at each monomer-addition step during the replication process: a reaction pathway involving several intermediate states associated with polymerases [10] or a reverse reaction catalyzed specifically by exonucleases [12]. Recently, proofreading based on a detailed polymerization mechanism coupled with cyclic pro-

tocols was proposed [46]. By contrast, our scheme is based on general thermodynamics and the multistep nature of template replication [47]. The error-correction effect works at each polymerization step, which is reinforced by the positive feedback from the template population, thus enabling long templates to avoid error catastrophe. This model can therefore serve as a guide for designing accurate template-replication systems and can further provide a plausible scenario for the transmission of sequence information in the prebiotic world.

APPENDIX A: REDUCTION OF THE MODEL INTO THE QS MODEL IN A CASE WITH SLOW DILUTION ($x_{0,p} \sim 0$)

We assume that polymerization is completed on a much faster time scale than that for the dilution (and the supply of primers) (i.e., $\phi \sim 0$, which is realized when $x_{0,p} \ll 1$). Because the last term is negligible in Eq. (2), using the steady-state condition $\dot{x}_{l,s} = 0$ allows adiabatic elimination of variables $x_{l,s}$, where $l \leqslant L - 1$, as

$$x_{l,s} = x_{l-1,s'} \frac{\sum_{S \in \{0,1\}^L} f(S)\beta(l-1,s',S)\nu(l,m,S)x_{L,S}}{(1+e^\Delta)\sum_{S \in \{0,1\}^L} f(S)\beta(l,s,S)x_{L,S}}. \tag{A1}$$

Because $s'm$ and s are identical sequences in Eq. (2), it follows that $\beta(l-1,s',S)\nu(l,m,S) = \beta(l,s,S)$. Therefore

$$x_{l,s} = x_{l-1,s'}/(1+e^\Delta). \tag{A2}$$

Using this equation, we can transform the rate equation for template polymers (i.e., polymers of length L) as follows:

$$\dot{x}_{L,s} = x_{L-1,s'} \sum_{S \in \{0,1\}^L} f(S)\beta(L-1,s',S)\nu(L,m,S)x_{L,S} - \phi x_{L,s}$$

$$= \frac{x_{0,p}}{(1+e^\Delta)^{L-1}} \sum_{S \in \{0,1\}^L} f(S)\beta(L-1,s',S)\nu(L,m,S)x_{L,S} - \phi x_{L,s}. \tag{A3}$$

Because $\beta(L-1,s',S)\nu(L,m,S) = \beta(L,s,S)$, it follows that

$$\dot{x}_{L,s} = \frac{x_{0,p}}{(1+e^\Delta)^{L-1}} \sum_{S \in \{0,1\}^L} f(S)\beta(L,s',S)x_{L,S} - \phi x_{L,s} = \frac{x_{0,p}}{(1+e^\Delta)^{L-1}} \sum_{S \in \{0,1\}^L} f(S)e^{\hat{n}(l,s,S)\Delta}x_{L,S} - \phi x_{L,s}$$

$$= x_{0,p}(1+e^\Delta) \sum_{S \in \{0,1\}^L} f(S)\frac{e^{(L-h_{s,s})\Delta}}{(1+e^\Delta)^L}x_{L,S} - \phi x_{L,s}, \tag{A4}$$

where $h_{i,j}$ is the Hamming distance between sequences i and j. Using $\mu = \frac{1}{1+e^\Delta}$, Eq. (4) in the main text is obtained.

APPENDIX B: DERIVATION OF THE UPPER LIMIT OF THE ACCURACY OF THE INFORMATION IN A CASE WITH FAST DILUTION ($x_{0,\mathrm{p}} \sim 1$)

First, we assume that the frequencies of "0" and "1" at different locations along the polymers are independent of each other. Let $\xi_{l,0}^{(i)}$ and $\xi_{l,1}^{(i)}$ denote the relative frequencies of polymers of sequence length l whose ith bit is "0" and "1," respectively, where $\xi_{l,0}^{(i)} + \xi_{l,1}^{(i)} = 1$. The concentration of polymers of sequence length l and sequence s is then expressed as follows:

$$x_{l,s} = x_l \prod_{i=1}^{l} \xi_{l,m_i}^{(i)}, \tag{B1}$$

where m_i is the ith bit of sequence s and x_l is the sum of the concentrations of the polymers of sequence length l ($x_l = \sum_{s \in \{0,1\}^l} x_{l,s}$).

From the steady state of Eq. (2), $x_{l,s}$ is calculated as follows:

$$x_{l,s} = x_{l-1,s'} \sum_{S \in \{0,1\}^L} r(l-1, s', m, S) x_{L,S} / \phi, \tag{B2}$$

where we assumed that the first and last terms are dominant in Eq. (2), because we assumed $x_{0,\mathrm{p}} \sim 1$ to allow a large ϕ. s' represents a sequence in which the end monomer of sequence s is deleted. Here, substituting Eq. (B1) and summing all the concentrations of the sequences whose ith bit is "0" gives

$$\phi \sum_{s \in s_0^{(i)}} x_{l,s} = \sum_{s \in s_0^{(i)}} x_{l-1} \prod_{j=1}^{l-1} \xi_{l-1,m_j}^{(j)} \sum_{S \in \{0,1\}^L} r(l-1, s', m_l, S) x_L \prod_{k=1}^{L} \xi_{L,M_k}^{(k)},$$

$$= x_{l-1} x_L \sum_{s \in s_0^{(i)}} \xi_{l-1,0}^{(i)} \prod_{j=1, j \neq i}^{l-1} \xi_{l-1,m_j}^{(j)} \sum_{S \in \{0,1\}^L} r(l-1, s', m_l, S) \prod_{k=1}^{L} \xi_{L,M_k}^{(k)}, \tag{B3}$$

where $\sum_{s \in s_0^{(i)}}$ denotes the summation of all of the sequences where the ith monomer is "0" and m_j and M_j denote the jth bits of the sequences s and S, respectively. Note that if $l = i$, we should read $\xi_{l-1,0}^{(i)}$ as $\xi_{l-1,0}^{(i)} = 1$. Here, we assume that $f(S) = f_0$ if S is the master sequence and $f(S) = f_1$ if S is the other sequence. This gives

$$\phi \sum_{s \in s_0^{(i)}} x_{l,s} = x_{l-1} x_L \sum_{s \in s_0^{(i)}} \xi_{l-1,0}^{(i)} \prod_{j=1, j \neq i}^{l-1} \xi_{l-1,m_j}^{(j)} \left(\sum_{S \in \{0,1\}^L} f_1 \nu(l, m_l, S) \beta(l-1, s', S) \prod_{k=1}^{L} \xi_{L,M_k}^{(k)} \right.$$

$$\left. + (f_0 - f_1) \nu(l, m_l, S_0) \beta(l-1, s', S_0) \prod_{k=1}^{L} \xi_{L,0}^{(k)} \right). \tag{B4}$$

By applying the definition of $\beta(l-1, s', S)$ and $\nu(l, m_l, S)$ for each pair of polymers, $x_{l,s} = x_l \prod_{i=1}^{l} \xi_{l,m_i}^{(i)}$ and a template

$$x_{L,S} = x_L \prod_{i=1}^{L} \xi_{l,M_i}^{(i)} = x_{l-1} x_L f_1 \xi_{l-1,0}^{(i)} \left(e^\Delta \xi_{L,0}^{(i)} + \xi_{L,1}^{(i)} \right) \prod_{j=1, j \neq i}^{l} \left(e^\Delta \xi_{l-1,0}^{(j)} \xi_{L,0}^{(j)} + \xi_{l-1,0}^{(j)} \xi_{L,1}^{(j)} + \xi_{l-1,1}^{(j)} \xi_{L,0}^{(j)} + e^\Delta \xi_{l-1,1}^{(j)} \xi_{L,1}^{(j)} \right)$$

$$+ x_{l-1} x_L (f_0 - f_1) e^\Delta \xi_{l-1,0}^{(i)} \prod_{k=1}^{L} \xi_{L,0}^{(k)} \prod_{j=1, j \neq i}^{l} \left(e^\Delta \xi_{l-1,0}^{(j)} + \xi_{l-1,1}^{(j)} \right), \tag{B5}$$

where we define $\xi_{l-1,0}^{(l)} = \xi_{l-1,1}^{(l)} = 1$. By using $\xi_{L,0}^{(i)} + \xi_{L,1}^{(i)} = 1$,

$$e^\Delta \xi_{l-1,0}^{(j)} + \xi_{l-1,1}^{(j)} = e^\Delta \xi_{l-1,0}^{(j)} \left(\xi_{L,0}^{(i)} + \xi_{L,1}^{(i)} \right) + \xi_{l-1,1}^{(j)} \left(\xi_{L,0}^{(i)} + \xi_{L,1}^{(i)} \right)$$

$$= e^\Delta \xi_{l-1,0}^{(j)} \xi_{L,0}^{(j)} + e^\Delta \xi_{l-1,0}^{(j)} \xi_{L,1}^{(j)} + \xi_{l-1,1}^{(j)} \xi_{L,0}^{(j)} + \xi_{l-1,1}^{(j)} \xi_{L,1}^{(j)}$$

$$= e^\Delta \xi_{l-1,0}^{(j)} \xi_{L,0}^{(j)} + e^\Delta \xi_{l-1,0}^{(j)} \xi_{L,1}^{(j)} + \xi_{l-1,1}^{(j)} \xi_{L,0}^{(j)} + \xi_{l-1,1}^{(j)} \xi_{L,1}^{(j)}$$

$$+ \xi_{l-1,0}^{(j)} \xi_{L,1}^{(j)} - \xi_{l-1,0}^{(j)} \xi_{L,1}^{(j)} + e^\Delta \xi_{l-1,1}^{(j)} \xi_{L,1}^{(j)} - e^\Delta \xi_{l-1,1}^{(j)} \xi_{L,1}^{(j)}$$

$$= \left(e^\Delta \xi_{l-1,0}^{(j)} \xi_{L,0}^{(j)} + \xi_{l-1,0}^{(j)} \xi_{L,1}^{(j)} + \xi_{l-1,1}^{(j)} \xi_{L,0}^{(j)} + e^\Delta \xi_{l-1,1}^{(j)} \xi_{L,1}^{(j)} \right) + (1 - e^\Delta) \left(\xi_{l-1,1}^{(j)} - \xi_{l-1,0}^{(j)} \right) \xi_{L,1}^{(j)}. \tag{B6}$$
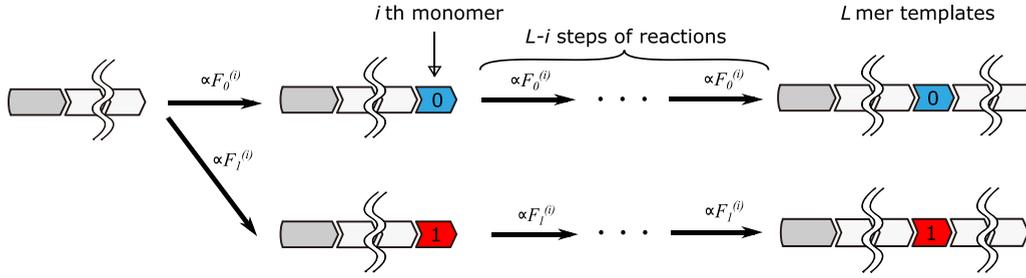
FIG. 5. Schematic of reaction pathways from the addition of the ith monomer to the completion of the template of length L whose ith monomer is "0" or "1." As in Fig. 1(a), each arrow represents the monomer addition to a substrate polymer using a template.

In the last line, we assume that the last term is much smaller than the first term, because $(1 - e^\Delta)(\xi_{l-1,1}^{(j)} - \xi_{l-1,0}^{(j)}) \sim 0$ if Δ is small and $\xi_{L,1}^{(j)} \sim 0$ if Δ is large; thus

$$e^\Delta \xi_{l-1,0}^{(j)} + \xi_{l-1,1}^{(j)} \sim \left(e^\Delta \xi_{l-1,0}^{(j)} \xi_{L,0}^{(j)} + \xi_{l-1,0}^{(j)} \xi_{L,1}^{(j)} + \xi_{l-1,1}^{(j)} \xi_{L,0}^{(j)} + e^\Delta \xi_{l-1,1}^{(j)} \xi_{L,1}^{(j)}\right). \tag{B7}$$

Similarly, we obtain the expressions for the sequences whose ith monomer is "1," $\sum_{s \in s_1^{(i)}} x_{l,s}$. Thus the relative production rate of a polymer with ith monomer "0" and "1" is given by

$$\phi \sum_{s \in s_0^{(i)}} x_{l,s} = \phi x_l \xi_{l,0}^{(i)} = A_l^{(i)} \left[e^\Delta \xi_0 f_0 + e^\Delta (\xi_0^{(i)} - \xi_0) f_1 + \xi_1^{(i)} f_1\right] x_{l-1} \xi_{l-1,0}^{(i)},$$

$$\phi \sum_{s \in s_1^{(i)}} x_{l,s} = \phi x_l \xi_{l,1}^{(i)} = A_l^{(i)} \left[\xi_0 f_0 + (\xi_0^{(i)} - \xi_0) f_1 + e^\Delta \xi_1^{(i)} f_1\right] x_{l-1} \xi_{l-1,1}^{(i)}, \tag{B8}$$

where $A_l^{(i)}$ is a constant that satisfies $A_l^{(i)} = x_L \prod_{j=1, j\neq i}^{l} (e^\Delta \xi_{l-1,0}^{(j)} + \xi_{l-1,1}^{(j)})$. Here, we define $F_0^{(i)}$ and $F_0^{(i)}$ as

$$F_0^{(i)} = e^\Delta \xi_0 f_0 + e^\Delta (\xi_0^{(i)} - \xi_0) f_1 + \xi_1^{(i)} f_1,$$
$$F_1^{(i)} = \xi_0 f_0 + (\xi_0^{(i)} - \xi_0) f_1 + e^\Delta \xi_1^{(i)} f_1, \tag{B9}$$

respectively, which are interpreted as the relative rates of the monomer addition to the sequence whose ith monomer is "0" and "1," respectively (Fig. 5). Note that $F_0^{(i)}$ and $F_0^{(i)}$ do not depend on the length of the sequence. Recursive application of this process allows the fraction of the template sequence with ith "0" or "1" monomer, $\xi_0^{(i)}$ or $\xi_1^{(i)}$, to be given by a self-consistent solution of

$$\frac{\xi_0^{(i)}}{\xi_1^{(i)}} = \frac{(F_0^{(i)})^{L+1-i}}{(F_1^{(i)})^{L+1-i}}, \tag{B10}$$

as explained in the main text. The numerical solution for Eq. (B10) in a case with $L = 4$ is plotted in Fig. 6.

APPENDIX C: THE ERROR THRESHOLD FOR A REPLICATING TEMPLATE WITH ERROR CORRECTION

The threshold value for the error catastrophe is roughly estimated in the case with a fast dilution limit (i.e., $x_{0,p} \sim 1$). Based on the discussion in the main text, for multi-step reactions, the effective error rate at the ith bit of the template during replication is modified as $\frac{\mu^{L+1-i}}{\mu^{L+1-i}+(1-\mu)^{L+1-i}}$ ($= \frac{1}{1+\exp(-(L+1-i)\Delta)}$). In this case, the error threshold at which the growth rate of the master sequence is overwhelmed by that of

the others is estimated from the condition

$$f_0 \prod_{i=1}^{L} \frac{1}{1 + e^{-i\Delta}} \sim f_1, \tag{C1}$$

where f_0 and f_1 are the fitness of the master sequence and the others, respectively. If we assume that L is infinitely large, then the threshold for μ is derived numerically as $\mu^* \sim 0.4268$ (Fig. 7). It should be noted that although the fraction of the master sequence ξ_0 is small if $\mu > \mu^*$, the threshold for
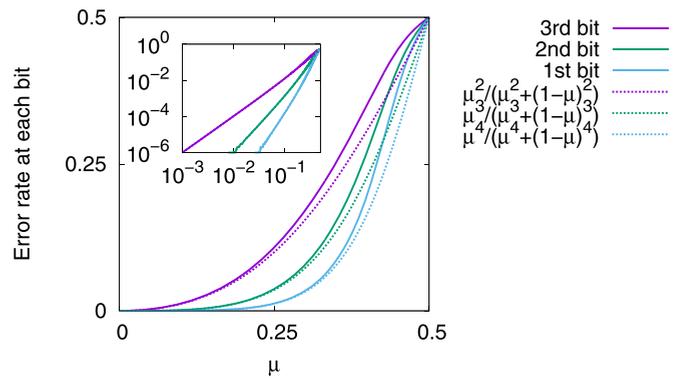
FIG. 6. The fraction of the error at each bit in the template sequence with length $L = 4$. The solid curves represent the fraction of the template polymer, $\xi_1^{(1)}$, $\xi_1^{(2)}$, and $\xi_1^{(3)}$. The dotted curves represent the minimum error rate that can be achieved in the KPR scheme of steps 2, 3, and 4, respectively. The inset shows the same plot with log-log axes.
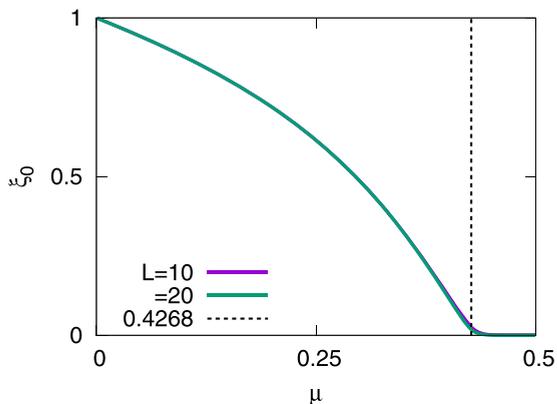
FIG. 7. The dependence on the error rate μ of the fraction of the master sequence among all of the templates ξ_0 in a case with $x_{0,p} \sim 1$. The dashed line represents the error-rate value μ^* in the solution of Eq. (C1) in a case with a large L limit.

$\mathcal{A}$ is higher than μ^*, because the difference between $\xi_0^{(i)}$ and $\xi_1^{(i)}$ is magnified exponentially.

APPENDIX D: RELAXATION DYNAMICS TOWARD THE STEADY STATE

In the main text, we discussed the steady state of the templates. Here, we discuss the relaxation dynamics toward reaching such a state.

In our model, a monomer incorporation reaction does not occur without a template. Hence long templates would not spontaneously appear if they are absent initially. Such templates are produced when including "spontaneous ligation reactions" from smaller monomers or polymers, as discussed previously [48,49]. Once this occurs, even if extremely rare, the same population of polymers and templates is reached, independent of their initial concentration.

Therefore we adopted the initial condition for the dynamics in which all template sequences exist uniformly in small amounts. We calculated the time course of the template distribution [Fig. 8(a)] and the accuracy $\mathcal{A}$ [defined in Eq. (3); Fig. 8(b)]. As expected, the dynamics eventually reached the steady state at which the master sequence is dominant, with $\mathcal{A} \sim 1$ (if the error rate is below the threshold). Notably, the

relaxation is slowed down under the fast dilution regime (i.e., large x_p). Accordingly, this suggests a trade-off between the "evolution speed" and the strength of the proofreading effect.

APPENDIX E: ENERGY FLUX TO DRIVE REPLICATION AND PROOFREADING

For kinetic proofreading [10] to work, energy influx is needed to drive the system toward the nonequilibrium condition; similarly, in our system, the supply of the primer and dilution corresponds to such energy influx. Then, the energy influx per template production is given as the inverse of the yield of the template, $1/x_L$ (recall that if the primer supply rate is ϕ and the total production rate of templates with length L is F, then $F = x_L \phi$ at the steady state).

As shown in Fig. 9, with an increase in the energy influx (per template production), $1/x_L$, the accuracy $\mathcal{A}$ increases until it saturates toward its maximum value.

APPENDIX F: CASES IN WHICH SHORTER POLYMERS CAN ALSO ACT AS TEMPLATES

In the main text, we assumed that only the longest polymers with length L act as templates. Here, we show that this assumption is not essential: Even if shorter polymers also act as templates, the maintenance of accuracy of the information is essentially preserved.

In this revised system, the addition of a monomer using templates with arbitrary lengths is represented as

$$X_{l,s} + X_{L',S} \rightarrow X_{l+1,sm} + X_{L',S},$$
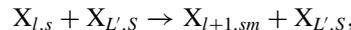
where $l \leqslant L' \leqslant L$. Then, we define the efficiency of sequence S with length L' as a template as $f(L', S)$. Here, we assume a single-peak fitness landscape: $f(L, S) \equiv f_0$ if $S = \{0\}^L$ (denoted by S_0), and $f(L', S) \equiv f_2$ if $1 \leqslant L' < L$; otherwise, $f(L, S) \equiv f_1 < f_0$.

We plotted the total concentration of sequences with length l, $x_l = \sum_{s \in \{0,1\}^l} x_{l,s}$ by varying the error rate μ (i.e., the binding energy Δ), as shown in Fig. 10(a). With the increase in μ (i.e., the decrease in Δ), the total concentration of the longest sequences x_L decreases, and the sequences with length L go extinct at a certain value of μ. Below this critical value of μ, the accuracy $\mathcal{A}$ of the information among the sequences with L [defined by Eq. (3)] does not significantly differ from that
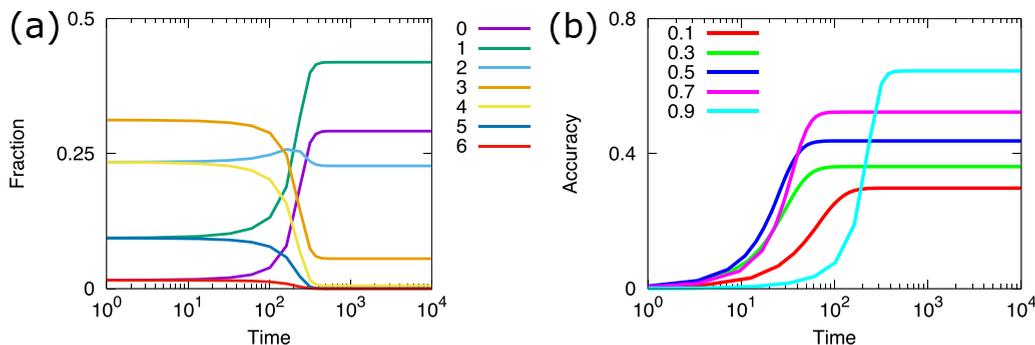
FIG. 8. (a) Time course of the fraction of sequences with the Hamming distance k from the master sequence S_0, ξ_k. We set the concentrations of the templates to be uniform ($x_{L,S} = 10^{-4}$ for all S) as the initial condition. We set $L = 6$, $\mu = 1/(1+e)$, and $x_{0,p} = 0.9$. (b) Time course of the accuracy $\mathcal{A}$. Each curve represents the difference in $x_{0,p}$.
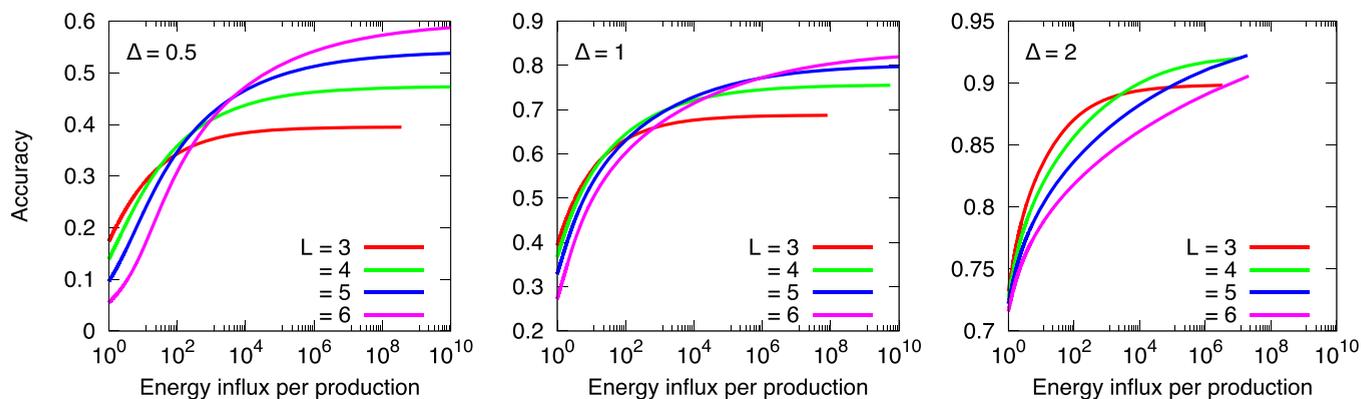
FIG. 9. Accuracy $\mathcal{A}$ [defined by Eq. (3) in the main text] vs the energy influx per template production $1/x_L$. We set $\Delta = 0.5$, 1, and 2.

found for the case in which only the sequences with L act as templates (i.e., $f_2 = 0$), as shown in Fig. 10(c). Note that even if the fitness for a shorter template f_2 is higher (i.e., $f_1 < f_2$), as long as $f_1 < f_0$, the results do not significantly differ, although the critical μ decreases, as shown in Figs. 10(b) and 10(d).

APPENDIX G: GENERAL FITNESS LANDSCAPES

In the main text, we considered only a single-peaked function as the fitness landscape $f(s)$. Here, we consider more complex functions as $f(s)$, which are discussed by Tarazona [4].
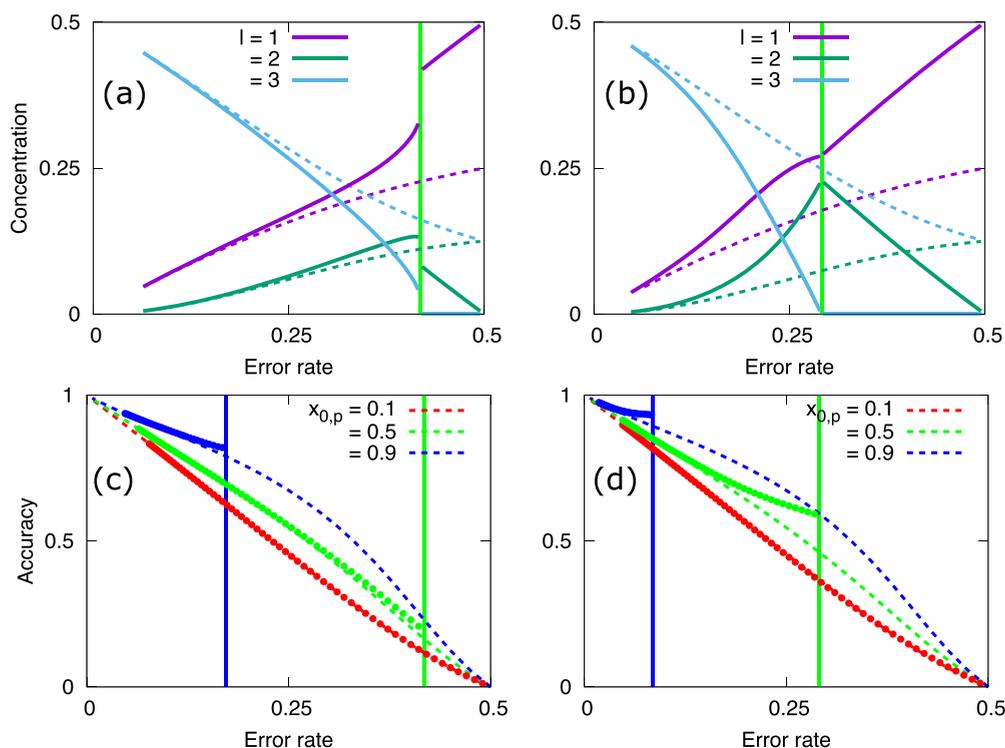
FIG. 10. (a) and (b) Total concentration of sequences with length l, plotted as a function of the error rate μ. We set $L = 3$, $x_{0,p} = 0.5$, and $f_2 = 0.1$ (a) or $f_2 = 1$ (b). The dashed curves represent the case in which only the longest sequences with L are templates (i.e., $f_2 = 0$), and the solid curves represent the results for the case in which shorter polymers with length l ($l < L$) also work as templates. In the latter case, the longest sequences with L go extinct at $\mu \sim 0.42$ in (a) and at $\mu \sim 0.29$ in (b), as indicated by vertical lines. (c) and (d) The dashed curves represent the accuracy $\mathcal{A}$ [defined as Eq. (3)] for the case in which only the longest sequences act as templates, varying the error rate μ, whereas dotted curves represent the results for the case in which shorter polymers also work as templates [$f_2 = 0.1$ in (c) and $f_2 = 1$ in (d)]. The difference in the colors represents the difference in the free primer concentration $x_{0,p}$ (i.e., the dilution rate ϕ). The vertical lines represent the error rate at which the longest sequences go extinct when $x_{0,p} = 0.5$ (green) and $x_{0,p} = 0.9$ (blue).
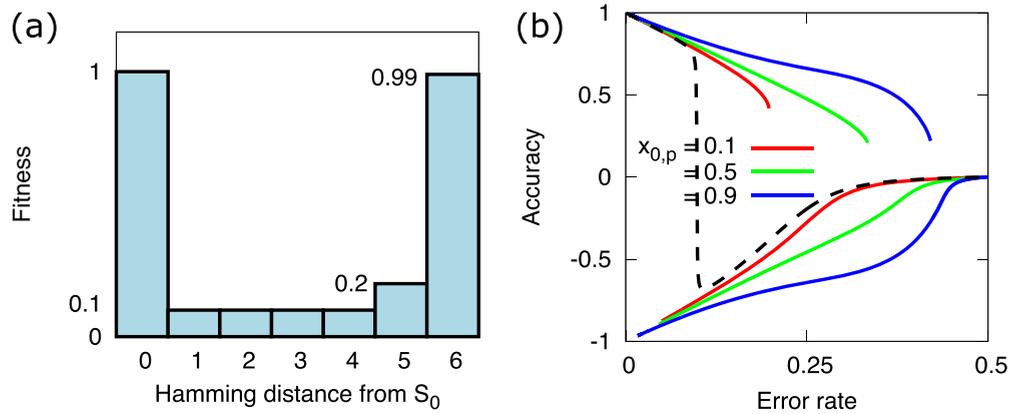
FIG. 11. (a) Fitness landscape for each sequence as a function of the Hamming distance from the master sequence S_0, defined by Eq. (G1) when $L = 6$. (b) The accuracy $\mathcal{A}$ [defined as Eq. (3)] calculated from the steady-state concentrations of templates under the quasidegenerate fitness landscape defined by Eq. (G1), plotted by varying the error rate μ. The difference in the colors represents the difference in the free primer concentration $x_{0,\mathrm{p}}$ (i.e., the dilution rate ϕ). The black dashed curve represents the result derived from the quasispecies model under the same fitness landscape. We set $L = 6$.

1. Quasidegenerate fitness landscape

As the first example, we investigate the following fitness function:

$$f(s) \equiv \begin{cases} f_0 & (h_{S_0,s} = 0) \\ f_L & (h_{S_0,s} = L) \\ f_{L-1} & (h_{S_0,s} = L - 1) \\ f_1 & (\text{otherwise}), \end{cases} \quad \text{(G1)}$$

where $h_{S_0,s}$ is the Hamming distance from the master sequence S_0.

If we set $f_0 = 1$, $f_1 = 0.1$, $f_L = 0.99$, and $f_{L-1} = 0.2$, this fitness landscape has two peaks: The sequence S_0 (all 0) has the highest fitness, whereas the sequence S_L (all 1) has the second-highest fitness with the local maximum [Fig. 11(a)].

Applying this fitness landscape to the QS model [Eq. (4) in the main text], there are three phases depending on the error

rate μ [Fig. 11(b)] [4]: For low μ, the sequence S_0 dominates (the accuracy $\mathcal{A} \sim 1$, as in the single-peak fitness landscape). A sudden jump occurs at a certain rate of μ, and the sequence S_L takes over as the dominant sequence ($\mathcal{A} \sim -1$). The next jump (i.e., the error catastrophe) then occurs at larger μ, at which the distribution of the sequences becomes uniform, and information is lost ($\mathcal{A} \sim 0$).

Similar phases appear by varying μ when applying this fitness landscape to our polymerization model [Eq. (2) in the main text]. However, for the intermediate μ value, the steady state is not monostable but rather is bistable [Fig. 11(b)]; that is, both the S_0-dominant ($\mathcal{A} \sim 1$) and the S_L-dominant ($\mathcal{A} \sim -1$) states are stable. Notably, the region of μ where the S_0-dominant state is stable expands as $x_{0,\mathrm{p}}$ increases. Due to the innate proofreading effect, the dominance of the (master) sequence S_0 is maintained even if the error rate is higher, and the error threshold is increased even under a two-peak fitness landscape.
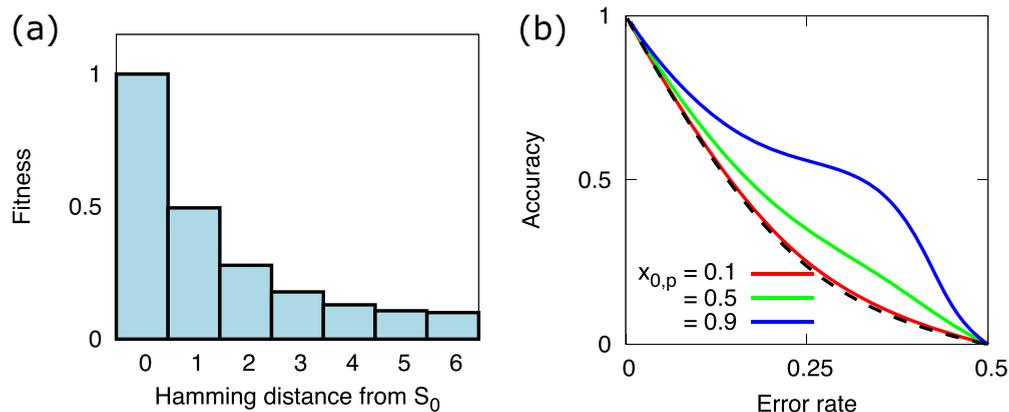
FIG. 12. (a) Fitness landscape for each sequence as a function of the Hamming distance from the master sequence S_0, defined as Eq. (G2) when $L = 6$. (b) The accuracy $\mathcal{A}$ [defined as Eq. (3)] calculated from the steady-state concentrations of templates under the smooth fitness landscape defined as Eq. (G2) by varying the error rate μ. The difference in the colors represents the difference in the free primer concentration $x_{0,\mathrm{p}}$ (i.e., the dilution rate ϕ). The black dashed curve represents the result derived from the quasispecies model under the same fitness landscape. We set $L = 6$, and $\exp(K/2) = 10$.
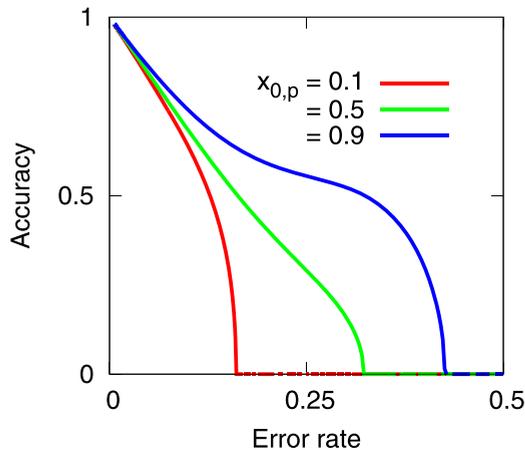
FIG. 13. The accuracy $\mathcal{A}$ [defined as Eq. (3)], which maintains the state of the dominance of the all-0 sequence, under the rugged fitness landscape defined as Eq. (G3) plotted as a function of the error rate μ. The colors represent the results for $x_{0,p}$, the free primer concentration (i.e., the dilution rate ϕ). We set $L = 6$. The sets of sequences with local maximum fitness $f(s)$ are given by $S^k = \{000000, 000101, 010101, 101010, 110100, 110110\}$.

2. Smooth fitness landscape

As the second example of an alternative fitness landscape, we consider

$$f(s) \equiv \exp\left(\frac{K}{2L^2} \sum_{i \neq j}^{L} (1 - m_i)(1 - m_j) \right)$$

$$= \exp\left(\frac{K}{2L^2} ((L - h_{S_0,s})^2 - L) \right), \qquad \text{(G2)}$$

where m_i represents the ith portion of sequence s. This fitness landscape is inspired by the ferromagnetic Ising model, where the sequences closer to homologous sequences (i.e., all

0) have higher fitness values [see Fig. 12(a) for the fitness landscape].

Compared with the general fitness landscape assumed in the model described in the main text, this fitness landscape is not sharp, although the master sequence S_0 still has the highest fitness. However, as shown in Fig. 12(b), the behavior of the accuracy when varying μ does not differ substantially from that under the sharp landscape [Fig. 2(b) in the main text].

3. Rugged fitness landscape

As the last example, we consider the rugged landscape derived from the Hopfield model for neural networks that embeds p patterns $S^k (= 1, \ldots, p)$ [50].

$$f(s) \equiv \exp\left(\frac{1}{2L^2} \sum_{k}^{p} K_k \sum_{i \neq j}^{L} \right.$$

$$\left. \times \left(1 - 2\hat{m}_i^k\right)\left(1 - 2\hat{m}_j^k\right)(1 - 2m_i)(1 - 2m_j) \right)$$

$$= \exp\left(\frac{1}{2L^2} \sum_{k}^{p} K_k ((L - 2h_{S^k,s})^2 - L) \right), \qquad \text{(G3)}$$

where $\hat{m}_i^k$ is the ith portion of the sequence S^k, which is randomly chosen among all sequences with length L. We set the master sequence $S^0 = S_0$ to have the highest fitness, and $p - 1$ sequences S^k also have local maximum fitness values. Recall that m_i is the ith portion of sequence s. K_k is the weight of sequence S^k: We set $K_k = 1$ if $k = 0$ and $K_k = \frac{1}{3}$ otherwise.

Under this rugged landscape, the accuracy $\mathcal{A}$ is plotted by varying the error rate μ in Fig. 13. Similar to the results obtained for the previous examples of fitness landscapes, the threshold value of μ increases with the increase in $x_{0,p}$. This threshold increase is observed independently of the specific choice of S^k.

[1] M. Eigen, Selforganization of matter and the evolution of biological macromolecules, Naturwissenschaften **58**, 465 (1971).

[2] J. Swetina and P. Schuster, Self-replication with errors: A model for polynucleotide replication, Biophys. Chem. **16**, 329 (1982).

[3] I. Leuthäusser, Statistical mechanics of Eigen's evolution model, J. Stat. Phys. **48**, 343 (1987).

[4] P. Tarazona, Error thresholds for molecular quasispecies as phase transitions: From simple landscapes to spin-glass models, Phys. Rev. A **45**, 6038 (1992).

[5] S. Franz and L. Peliti, Error threshold in simple landscapes, J. Phys. A: Math. Gen. **30**, 4481 (1997).

[6] D. B. Saakian and C.-K. Hu, Exact solution of the Eigen model with general fitness functions and degradation rates, Proc. Natl. Acad. Sci. USA **103**, 4935 (2006).

[7] N. Wagner, E. Tannenbaum, and G. Ashkenasy, Second Order Catalytic Quasispecies Yields Discontinuous Mean Fitness at Error Threshold, Phys. Rev. Lett. **104**, 188101 (2010).

[8] N. Takeuchi and P. Hogeweg, Evolutionary dynamics of RNA-like replicator systems: A bioinformatic approach to the origin of life, Phys. Life Rev. **9**, 219 (2012).

[9] E. Domingo, P. Schuster, and M. B. Oldstone, *Quasispecies: From Theory to Experimental Systems*, Current Topics in Microbiology and Immunology Vol. 392 (Springer, New York, 2016).

[10] J. J. Hopfield, Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity, Proc. Natl. Acad. Sci. USA **71**, 4135 (1974).

[11] J. Ninio, Kinetic amplification of enzyme discrimination, Biochimie **57**, 587 (1975).

[12] C. H. Bennett, Dissipation-error tradeoff in proofreading, Biosystems **11**, 85 (1979).

[13] A. Murugan, D. A. Huse, and S. Leibler, Speed, dissipation, and error in kinetic proofreading, Proc. Natl. Acad. Sci. USA **109**, 12034 (2012).

[14] P. Sartori and S. Pigolotti, Kinetic versus Energetic Discrimination in Biological Copying, Phys. Rev. Lett. **110**, 188101 (2013).

[15] S. Pigolotti and P. Sartori, Protocols for copying and proofreading in template-assisted polymerization, J. Stat. Phys. **162**, 1167 (2016).

[16] W. D. Piñeros and T. Tlusty, Kinetic proofreading and the limits of thermodynamic uncertainty, Phys. Rev. E **101**, 022415 (2020).

[17] V. Galstyan, K. Husain, F. Xiao, A. Murugan, and R. Phillips, Proofreading through spatial gradients, eLife **9**, e60415 (2020).

[18] D. Andrieux and P. Gaspard, Nonequilibrium generation of information in copolymerization processes, Proc. Natl. Acad. Sci. USA **105**, 9516 (2008).

[19] T. E. Ouldridge and P. R. ten Wolde, Fundamental Costs in the Production and Destruction of Persistent Polymer Copies, Phys. Rev. Lett. **118**, 158103 (2017).

[20] A primer is interpreted as a specific short polymer sequence. Generally, primers are required to initiate the polymerization reaction in experimental template-directed replication, given that a monomer alone cannot bind to a template polymer.

[21] Complementarity and directionality of templates are ignored for simplicity.

[22] The Boltzmann constant multiplied by the temperature, $k_B T$, is considered to be unity.

[23] In principle, E_0 can be sufficiently large considering appropriate physicochemical conditions, such as the diluted (low concentration) polymers or external driving force to separate the bond between polymers.

[24] This condition sets the upper limit of the differences in the reaction rates due to the difference in the binding energy of pairings.

[25] Here, we assume that the joining of a monomer to a polymer is an irreversible reaction. This situation is realized if an energetically activated monomer is used so that covalent bond formation is energetically favorable.

[26] By substituting the condition at the steady state $\sum_{l=1}^{L} \sum_{s \in \{0,1\}^l} \dot{x}_{l,s} = 0$ for Eq. (2), ϕ is determined as $x_{0,p}(1 + e^{\Delta}) \sum_{S \in \{0,1\}^L} f(S) x_{L,S} - (1 - x_{0,p})\phi = 0$. Note that we assume the boundary condition so that $x_{0,p}$ is a constant.

[27] The relaxation dynamics toward the steady state are discussed in Appendix D.

[28] This situation is satisfied, at least under the fitness landscape assumed in this paper.

[29] This monotonic increase in the accuracy $\mathcal{A}$ is observed only if the proofreading scheme is fully utilized, assuming both the infinite dilution rate and no saturation of bounded polymers.

[30] This condition is achieved by choosing E_0 such that $\hat{E}_a \gg 0$, where $\hat{E}_a$ is the energy barrier for a template to bind to the other template, $\hat{E}_a = E_0 - L\Delta$.

[31] Z. Varga and E. Szathmáry, An extremum principle for parabolic competition, Bull. Math. Biol. **59**, 1145 (1997).

[32] P. R. Wills, S. A. Kauffman, B. M. Stadler, and P. F. Stadler, Selection dynamics in autocatalytic systems: Templates replicating through binary ligation, Bull. Math. Biol. **60**, 1073 (1998).

[33] We shall discuss polymerization- and template-replication dynamics with product inhibition in a separate manuscript, which is in preparation.

[34] C. Fernando, G. Von Kiedrowski, and E. Szathmáry, A stochastic model of nonenzymatic nucleic acid replication:

"elongators" sequester replicators, J. Mol. Evol. **64**, 572 (2007).

[35] B. Obermayer, H. Krammer, D. Braun, and U. Gerland, Emergence of Information Transmission in a Prebiotic RNA Reactor, Phys. Rev. Lett. **107**, 018101 (2011).

[36] A. V. Tkachenko and S. Maslov, Spontaneous emergence of autocatalytic information-coding polymers, J. Chem. Phys. **143**, 045102 (2015).

[37] S. Toyabe and D. Braun, Cooperative Ligation Breaks Sequence Symmetry and Stabilizes Early Molecular Replication, Phys. Rev. X **9**, 011056 (2019).

[38] In the QS model, if shorter polymers of length L' could work as well as the long templates, the shorter templates will replicate faster and thus outcompete the longer ones. In this case, the length L of a template can be replaced by a smaller L', representing the so-called Spiegelman's monster problem [51].

[39] M. Kreysing, L. Keil, S. Lanzmich, and D. Braun, Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length, Nat. Chem. **7**, 203 (2015).

[40] R. Mizuuchi, A. Blokhuis, L. Vincent, P. Nghe, N. Lehman, and D. Baum, Mineral surfaces select for longer RNA molecules, Chem. Commun. **55**, 2090 (2019).

[41] Note that in multimodal landscapes, the system could show multistability because of frequency-dependent selection among templates [37,49,52].

[42] S. Rajamani, J. K. Ichida, T. Antal, D. A. Treco, K. Leu, M. A. Nowak, J. W. Szostak, and I. A. Chen, Effect of stalling after mismatches on the error catastrophe in nonenzymatic nucleic acid replication, J. Am. Chem. Soc. **132**, 5880 (2010).

[43] N. Prywes, J. C. Blain, F. Del Frate, and J. W. Szostak, Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated oligonucleotides, eLife **5**, e17756 (2016).

[44] G. F. Joyce and J. W. Szostak, Protocells and RNA self-replication, Cold Spring Harbor Perspect. Biol. **10**, a034801 (2018).

[45] K. Le Vay, L. I. Weise, K. Libicher, J. Mascarenhas, and H. Mutschler, Templated self-replication in biomimetic systems, Adv. Biosys. **3**, 1800313 (2019).

[46] T. Göppel, B. Obermayer, I. A. Chen, and U. Gerland, A kinetic error filtering mechanism for enzyme-free copying of nucleic acid sequences, bioRxiv.

[47] With respect to multistep polymerization, the influence of a single mismatch can be magnified through the polymerization RecA protein in the homology search [53–55].

[48] Y. J. Matsubara and K. Kaneko, Optimal size for emergence of self-replicating polymer system, Phys. Rev. E **93**, 032503 (2016).

[49] Y. J. Matsubara and K. Kaneko, Kinetic Selection of Template Polymer with Complex Sequences, Phys. Rev. Lett. **121**, 118101 (2018).

[50] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).

[51] S. Spiegelman, I. Haruna, I. Holland, G. Beaudreau, and D. Mills, The synthesis of a self-propagating and infectious nucleic acid with a purified enzyme, Proc. Natl. Acad. Sci. USA **54**, 919 (1965).

[52] P. W. Anderson, Suggested model for prebiotic evolution: The use of chaos, Proc. Natl. Acad. Sci. USA **80**, 3386 (1983).

[53] R. Bar-Ziv, T. Tlusty, and A. Libchaber, Protein–DNA computation by stochastic assembly cascade, Proc. Natl. Acad. Sci. USA **99**, 11589 (2002).

[54] T. Tlusty, R. Bar-Ziv, and A. Libchaber, High-Fidelity DNA Sensing by Protein Binding Fluctuations, Phys. Rev. Lett. **93**, 258103 (2004).

[55] D. Sagi, T. Tlusty, and J. Stavans, High fidelity of RecA-catalyzed recombination: A watchdog of genetic diversity, Nucleic Acids Res. **34**, 5021 (2006).