Supporting Information: Thermodynamic Cost, Speed, Fluctuations, and Error Reduction of Biological Copy Machines

Yonghyun Song[1] and Changbong Hyeon[1]
[1]*Korea Institute for Advanced Study, Seoul 02455, Korea*

CONTENTS

BENNETT MODEL

In the non-proofreading model of copying processes introduced by Bennett [1], a copying enzyme synthesizes the complementary polymer strand by incorporating the monomers from the solution via a single kinetic step (Fig. S1A). The copy process is described by three key parameters: the binding free energies of correct and incorrect monomers, $\Delta\mu_c$ and $\Delta\mu_i$, and the difference in the kinetic barriers, δ. Here, we keep the difference between the binding free energies ($\Delta\Delta\mu_{ci} \equiv \Delta\mu_c - \Delta\mu_i$) and the difference between the kinetic barriers constant. For instance, in DNA replication, $\Delta\Delta\mu_{ci}$ and δ are determined by the molecular properties of the nucleotides and the polymerase. The error probability can be modulated by changing nucleotide concentrations, which corresponds to changing $\Delta\mu_c$ in the Bennett model. Thus, in the Bennett model, η is evaluated as a function of $\Delta\mu_c$, and vice versa. In the following, we derive the expressions of $\Delta\mu_c(\eta)$, $\lambda(\eta)$, $\mathcal{A}(\eta)$ and $Q(\eta)$. Using these expressions, we plot the diagrams of η and Q as functions of δ and $\Delta\mu_c$ (Fig. S1B-E).

The evolution of the probability P of the complementary polymer can be described by the following master equation,

$$\begin{aligned}\dot{P}(\dots c) =& k_f^c P(\dots) + k_r^c P(\dots cc) + k_r^i P(\dots ci) \\ & - (k_f^c + k_f^i + k_r^c) P(\dots c), \\ \dot{P}(\dots i) =& k_f^i P(\dots) + k_r^c P(\dots ic) + k_r^i P(\dots ii) \\ & - (k_f^c + k_f^i + k_r^i) P(\dots i),\end{aligned} \tag{S1}$$

where the sequence of the complementary polymer is represented by $(\dots)$, $(\dots c)$, $(\dots i)$, and so on as in Fig. S1A.

Assuming that the error probability of each position of the copy polymer is independent from the prior sequence, we can make the following substitutions: $P(\dots c) = P(\dots)P(c)$, $P(\dots ci) = P(\dots)P(c)P(i)$, and so forth. Then, at steady state, eq. S1 takes the form

$$\begin{aligned} k_f^c + k_r^c P(c)^2 + k_r^i P(c)P(i) - (k_f^c + k_f^i + k_r^c) P(c) = 0 \\ k_f^i + k_r^c P(c)P(i) + k_r^i P(i)^2 - (k_f^c + k_f^i + k_r^i) P(i) = 0. \end{aligned}$$

By rearranging the above equations, we can write the error probability, η, as

$$\eta = P(i) = \frac{\langle J^i \rangle}{\langle J^i \rangle + \langle J^c \rangle} = \frac{k_f^i - \eta k_r^i}{k_f^i - \eta k_r^i + k_f^c - (1-\eta) k_r^c}, \tag{S2}$$

where $\langle J^c \rangle$ and $\langle J^i \rangle$ are defined as the average reaction currents for correct and incorrect monomers, respectively,

$$\begin{aligned}\langle J^c \rangle &= k_f^c - (1-\eta) k_r^c \\ \langle J^i \rangle &= k_f^i - \eta k_r^i.\end{aligned} \tag{S3}$$

Essentially, we have transformed the dynamics along the tree structure (Fig S1A) into a Markov process. The equivalence of the expressions of η from eq. S1 and eq. S2 has been shown for the double-cyclic reversible 3-state nework model in ref. [2]. More general treatment of the dynamics that occur along tree structures can be found in refs. [3–5]. In the present work, the main conclusions pertaining to η (and Q) are further supported by explicit simulations of the master equation representation (Fig. 2B,C and Fig. 4C,D)

At the detailed balance (DB) condition, no current should flow through both pathways associated with correct and incorrect monomer incorporations, i.e., $\langle J^c \rangle = 0$ and $\langle J^i \rangle = 0$; $\beta\mathcal{A}^c = \ln\left(\frac{k_f^c}{(1-\eta)k_r^c}\right) = 0$ and $\beta\mathcal{A}^i = \ln\left(\frac{k_f^i}{\eta k_r^i}\right) = 0$, which leads to

$$1 - \eta_{eq} = \left(\frac{k_f^c}{k_r^c}\right)_{eq} \equiv e^{-\beta\Delta\mu_{c,eq}} \tag{S4}$$

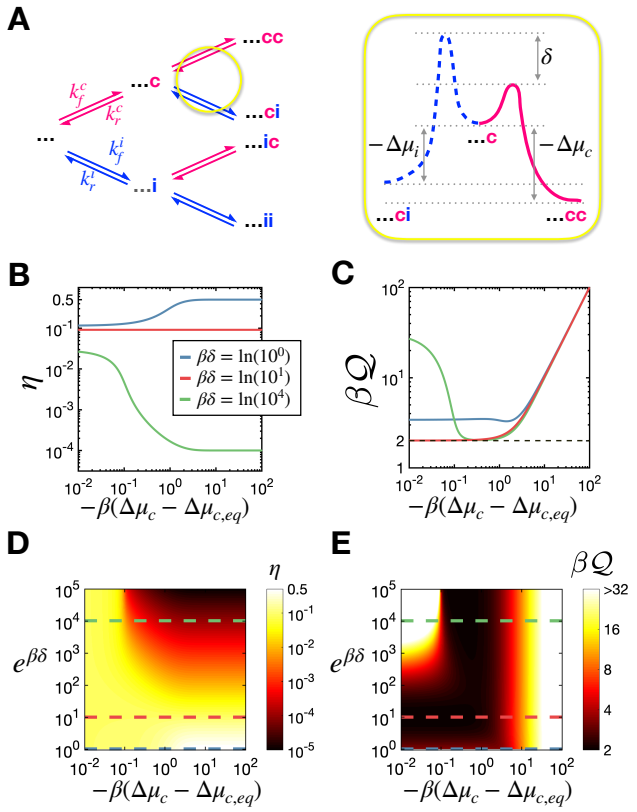

FIG. S1. The reaction kinetics of a non-proofreading copying system generalizing Bennett model. (A) (Left) The schematic of the copy polymer elongation. The (...) represents the previously synthesized copy polymer. (Right) The schematic of the free energy profile of monomer incorporation. The free energies of incorporating correct and incorrect monomers are $\beta\Delta\mu_c = -\ln(k_f^c/k_r^c)$ and $\beta\Delta\mu_i = -\ln(k_f^i/k_r^i)$, respectively. The incorporation of the correct monomer occurs faster than that of the incorrect monomer, as described by the parameter $\beta\delta = \ln(k_f^c/k_f^i) > 0$. (B) The error probability (η) and (C) Q as functions of the chemical potential bias for different values of the kinetic discrimination parameter ($\beta\delta$), where $e^{-\beta(\Delta\mu_c-\Delta\mu_i)} = 10$. The curves in (C) for various $\beta\delta$ values are color-coded identically as those in (B). The dashed curves in (D) and (E) are also color-coded identically as those in (B). In (C), the physical limit of $\beta Q = 2$ is marked with a dashed line. Diagrams of (D) η and (E) βQ as a function of $e^{\beta\delta}$ and $-\beta(\Delta\mu_c - \Delta\mu_{c,eq})$.

and

$$\eta_{eq} = \left(\frac{k_f^i}{k_r^i}\right)_{eq} \equiv e^{-\beta\Delta\mu_{i,eq}} \tag{S5}$$

We note that in order for η_{eq} to be in the range of $0 < \eta_{eq} < 1$, $\beta\Delta\mu_{c,eq}$ and $\beta\Delta\mu_{i,eq}$ should be positive, meaning that chemical potential bias of monomer incorporation is positive (uphill).

Then, by taking the ratio between eqs. S4 and S5, we obtain

$$\eta_{eq} = \frac{1}{1 + e^{-\beta\Delta\Delta\mu_{ci}}}. \tag{S6}$$

where $\Delta\Delta\mu_{ci} = \Delta\mu_{c,eq} - \Delta\mu_{i,eq}$. At the limit of strongly forward driven reactions, i.e., $k_f^c \gg (1-\eta)k_r^c$ and $k_f^i \gg \eta k_r^i$, eq. S2 is led to

$$\eta \to k_f^i/(k_f^i + k_f^c) = 1/(1 + e^{\beta\delta}) \equiv \eta_0. \tag{S7}$$

Next, we can write eq. S2 as

$$\begin{aligned}\eta &= \frac{k_f^i/k_r^i - \eta}{k_f^i/k_r^i - \eta + (k_r^c/k_r^i)\left(k_f^c/k_r^c - (1-\eta)\right)} \\ &= \frac{e^{-\beta\Delta\mu_i} - \eta}{(e^{-\beta\Delta\mu_i} - \eta) + e^{-\beta(\Delta\Delta\mu_{ci}-\delta)}\left(e^{-\beta\Delta\mu_c} - (1-\eta)\right)}\end{aligned} \tag{S8}$$

where $k_r^c/k_r^i = e^{-\beta(\Delta\Delta\mu_{ci}-\delta)}$ was used. After some rearrangements, we can express $\Delta\mu_c$ as a function of η, η_0, and η_{eq} as follows

$$-\beta\Delta\mu_c = \ln\frac{\eta(1-\eta)(\eta_{eq}-\eta_0)}{(\eta-\eta_0)\eta_{eq}}. \tag{S9}$$

Next, $\mathcal{A}$ can be written as

$$\begin{aligned}\beta\mathcal{A}(\eta) &= \frac{1}{\langle J^c\rangle + \langle J^i\rangle}\left[\langle J^c\rangle \ln\frac{k_f^c}{(1-\eta)k_r^c} + \langle J^i\rangle \ln\frac{k_f^i}{\eta k_r^i}\right] \\ &= (1-\eta)\ln\frac{k_f^c}{(1-\eta)k_r^c} + \eta\ln\frac{k_f^i}{\eta k_r^i} \\ &= \underbrace{-\beta\left[(1-\eta)\Delta\mu_c + \eta\Delta\mu_i\right]}_{=-\beta\Delta\mu} + \underbrace{\left[-(1-\eta)\ln(1-\eta) - \eta\ln\eta\right]}_{=I} \\ &= \eta\ln\frac{\eta_{eq}(1-\eta)}{\eta(1-\eta_{eq})} + \ln\frac{\eta(\eta_{eq}-\eta_0)}{\eta_{eq}(\eta-\eta_0)}.\end{aligned} \tag{S10}$$

As expected, $\lim_{\eta\to\eta_{eq}}\beta\mathcal{A}(\eta) = 0$ and $\lim_{\eta\to\eta_0}\beta\mathcal{A}(\eta) = \infty$, which means that η approaches η_{eq} and η_0 at the zero and infinite dissipation limits, respectively.

Importantly, $\mathcal{A}$ can be decomposed into two contributions (eq. S10): $-\beta\Delta\mu$ is the free energy gain after the monomer incorporation, and I is the Shannon information entropy arising from the chance of incorporating correct (c) and incorrect (i) monomers to the copy strand. The information (I) is maximized to $I = \ln 2$ when the odds of incorporating the correct and incorrect monomers is identical ($\eta = 1/2$), whereas $I = 0$ if only the correct or incorrect monomers are incorporated. This implies that as long as the chemical potential of monomers in solution is constantly maintained, the process near the DB condition ($\beta\mathcal{A} = -\beta\Delta\mu + I \gtrsim 0$) can still be driven by the entropy $I(\geq \beta\Delta\mu)$ even if the polymerization is energetically uphill ($\Delta\mu > 0$) [1].

The Fano factor λ can be calculated as

$$\begin{aligned}\lambda(\eta) &= \frac{\langle(\delta J^c)^2\rangle + \langle(\delta J^i)^2\rangle}{\langle J^c\rangle + \langle J^i\rangle} \\ &= \frac{(k_f^i + \eta k_r^i) + (k_f^c + (1-\eta)k_r^c)}{(k_f^i - \eta k_r^i) + (k_f^c - (1-\eta)k_r^c)} \\ &= \frac{\left(k_f^i/k_r^i + \eta\right) + (k_r^c/k_r^i)\left(k_f^c/k_r^c + (1-\eta)\right)}{\left(k_f^i/k_r^i - \eta\right) + (k_r^c/k_r^i)\left(k_f^c/k_r^c - (1-\eta)\right)} \\ &= \frac{\left(e^{-\beta\Delta\mu_i} + \eta\right) + e^{-\beta(\Delta\Delta\mu_{ci}-\delta)}\left(e^{-\beta\Delta\mu_c} + (1-\eta)\right)}{\left(e^{-\beta\Delta\mu_i} - \eta\right) + e^{-\beta(\Delta\Delta\mu_{ci}-\delta)}\left(e^{-\beta\Delta\mu_c} - (1-\eta)\right)} \\ &= \frac{2(\eta_{eq}-\eta_0)\eta^2 + (\eta_0 + \eta_0^2 - 2\eta_{eq})\eta + (1-\eta_0)\eta_0\eta_{eq}}{\eta_0(1-\eta_0)(\eta - \eta_{eq})}.\end{aligned} \tag{S11}$$

$\mathcal{Q}(\eta)$ evaluated using the expression of $\mathcal{A}(\eta)$ and $\lambda(\eta)$, i.e., $\mathcal{Q}(\eta) = \mathcal{A}(\eta)\lambda(\eta)$, quantifies the translational efficiency of the copying enzyme along the template polymer [6, 7]. For strongly driven systems ($\mathcal{A} \gg 0$), all the curves of $\mathcal{Q}(\mathcal{A})$ with different values of δ converge (Fig. S1C). However, near the DB condition, where I contributes significantly to $\mathcal{A}$, $\mathcal{Q}$ shows complex dependence on δ.

At the DB condition,

$$\beta\mathcal{Q}(\eta_{eq}) = 2 + \frac{(\eta_{eq}-\eta_0)^2}{(1-\eta_0)\eta_0} \geq 2. \tag{S12}$$

Thus, the lower bound $2\ k_BT$ is attained at the DB condition when $\eta_{eq} = \eta_0$.

$\mathcal{Q}$ can also approach its lower bound $2\ k_BT$ at the limiting condition of $\beta\delta \gg 1$. At this limit, only correct monomers are incorporated into the copy polymer ($\langle J^i\rangle = 0$), which leads to

$$\mathcal{Q} = \mathcal{A}\frac{\langle(\delta J^c)^2\rangle}{\langle J^c\rangle} = \mathcal{A}\frac{k_f^c + k_r^c}{k_f^c - k_r^c} = \mathcal{A}\frac{e^{\beta\mathcal{A}}+1}{e^{\beta\mathcal{A}}-1} \geq 2\ k_BT, \tag{S13}$$

and $\lim_{\mathcal{A}\to 0,\delta\to\infty}\mathcal{Q} = 2\ k_BT$. The two limiting scenarios at which $\mathcal{Q}$ approaches $2\ k_BT$ can be seen in Fig. S1E.

MICHAELIS-MENTEN REACTIONS

We provide conditions at which $\mathcal{Q}$ has a local minimum with respect to substrate concentration ($[S]$) in reversible MM reactions shown in Fig. S2A. First, the affinity ($\mathcal{A}$),

$$\beta\mathcal{A}([S]) = \ln\left(\frac{k_{on}k_{cat}[S]}{k_{off}k_{rev}[P]}\right), \tag{S14}$$

is a strictly increasing function of [S]. Next, the Fano factor (λ) as a function of [S] is given by

$$\lambda([S]) = \frac{k_{on}k_{cat}[S] + k_{off}k_{rev}[P] - 2\left(\frac{k_{on}k_{cat}[S]-k_{off}k_{rev}[P]}{k_{on}[S]+k_{off}+k_{cat}+k_{rev}[P]}\right)^2}{k_{on}k_{cat}[S] - k_{off}k_{rev}[P]} \tag{S15}$$

$$= \frac{e^{\beta\mathcal{A}}+1}{e^{\beta\mathcal{A}}-1} - 2\frac{\gamma^2\left(e^{\beta\mathcal{A}}-1\right)}{\left(\gamma^2+\gamma\xi+e^{\beta\mathcal{A}}\right)^2}, \tag{S16}$$

where $\gamma = \frac{k_{cat}}{\sqrt{k_{off}k_{rev}[P]}}$ and $\xi = \frac{k_{off}+k_{rev}[P]}{\sqrt{k_{off}k_{rev}[P]}}$ are dimensionless constants. λ has a local minimum only when $\gamma > 1$, or equivalently, when $k_{cat} > \sqrt{k_{off}k_{rev}[P]}$. At the limit of a strongly driven catalytic step ($k_{cat} \gg k_{rev}[P]$), the expression for λ simplifies to

$$\lambda([S]) \approx 1 - \frac{2k_{cat}[S]}{k_{on}([S]+K_m)^2}, \tag{S17}$$

which is minimized at $[S] = K_m$ with $K_m = (k_{off}+k_{cat})/k_{on}$ (Fig. S2B).

Since $\mathcal{A}$ is monotonic with [S], a local minimum of $\mathcal{Q}([S]) = \mathcal{A}([S])\lambda([S])$ can only occur near to that of $\lambda([S])$. Using this, we numerically determined the range of γ and ξ values at which $\mathcal{Q}([S])$ has a local minimum away from the DB condition. For any ξ, there exists a γ above which $\mathcal{Q}[S]$ is non-monotonic (Fig. S2B). Thus, when $k_{cat} \gg k_{rev}[P]$ and k_{cat} is sufficiently larger than $\sqrt{k_{off}k_{rev}[P]}$, $\mathcal{Q}([S])$ has a local minimum around $[S] \approx K_m$.

MATHEMATICAL EXPRESSIONS OF η AND $\mathcal{A}$

Here, we provide the details of obtaining η and $\mathcal{Q}$ in the main text. A more mathematically rigorous treatment on the subject can be found in ref. [8].

In the following, without loss of generality, we will define the error probability (η) and affinity ($\mathcal{A}$) of copy processes by referring to the kinetic mechanism of mRNA translation. To begin, consider the ribosome at position l of the mRNA sequence, decoding a specific codon type. For each amino acid type a, there exists a set T_a^l, of associated aa-tRNAs, each of which represents a separate kinetic path of incorporating a into the protein. If the codon at position l is GGG, amino acids Gly, Ala, Arg, Glu, Trp, and Val can be polymerized, with the following set of associated tRNAs: $T_{\text{Gly}}^l = \{\text{tRNA}_1^{\text{Gly}},\ \text{tRNA}_2^{\text{Gly}},\ \text{tRNA}_3^{\text{Gly}}\}$, $T_{\text{Ala}}^l = \{\text{tRNA}_{1\text{B}}^{\text{Ala}}\}$, $T_{\text{Arg}}^l = \{\text{tRNA}_3^{\text{Arg}},\ \text{tRNA}_5^{\text{Arg}}\}$, $T_{\text{Glu}}^l = \{\text{tRNA}_2^{\text{Glu}}\}$, $T_{\text{Trp}}^l = \{\text{tRNA}^{\text{Trp}}\}$, and $T_{\text{Val}}^l = \{\text{tRNA}_1^{\text{Val}}\}$ (Fig. S7).

At steady state, we assume that, η_a^l, the probability of incorporating amino acid a at position l, where the codon is specified, is given by

$$\eta_a^l = \frac{\sum_{t\in T_a^l}\langle J_{a,t}^{l,\text{pol}}\rangle}{\sum_{\alpha\in\{\text{aa}\}}\sum_{t\in T_\alpha^l}\langle J_{\alpha,t}^{l,\text{pol}}\rangle}, \tag{S18}$$

where {aa} is the set of all amino acids, and $\langle J_{a,t}^{l,\text{pol}}\rangle$ is the polymerization current of aa-tRNA t.

Next, we define the affinity associated with polymerization using the previously discussed example of Gly incorporation at codon GGG. The polymerization affinity of Gly along the cognate kinetic path associated with Gly-$\text{tRNA}_1^{\text{Gly}}$ is

$$\beta\mathcal{A}_{\text{Gly,Gly1}}^{l,\text{pol}} = -\beta\Delta\mu_{\text{Gly,Gly1}}^{l,\text{pol}} - \ln(\eta_{\text{Gly}}^l), \tag{S19}$$

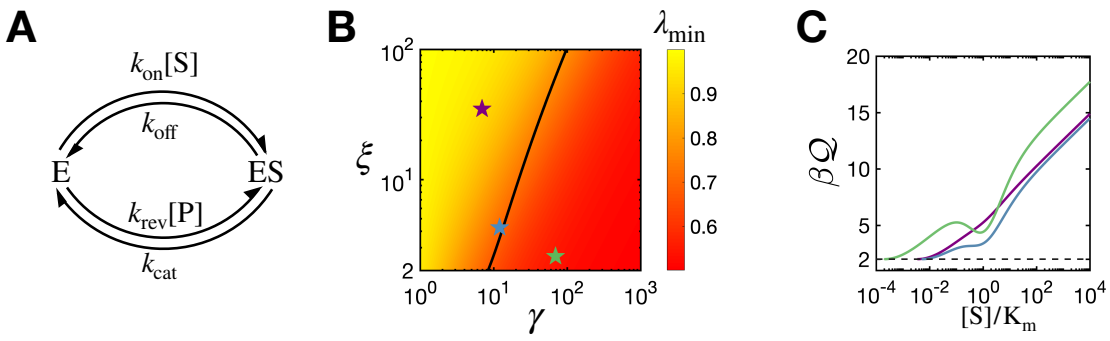

FIG. S2. Michaelis-menten type enzyme kinetics. (A) Schematic of MM type enzyme kinetics. (B) The minimum Fano factor with respect to [S] ($\lambda_{\min}$) plotted against dimensionless constants $\gamma = \frac{k_{cat}}{\sqrt{k_{off} k_{rev}[P]}}$ and $\xi = \frac{k_{off}+k_{rev}[P]}{\sqrt{k_{off} k_{rev}[P]}}$. To the left (resp. right) of the black curve, Q is monotonic (resp. non-monotonic). (C) Q plotted as a function of [S] normalized by $K_m = \frac{k_{off}+k_{cat}}{k_{on}}$. The corresponding γ and ξ values are shown in (B) by the star symbols with the matching color. The dotted line demarcates the lower bound $Q = 2k_BT$. The kinetic rate constants used for the plots are as follows: Green: $k_{on} = 10^8\ M^{-1}s^{-1}$, $k_{off} = 30\ s^{-1}$, $k_{cat} = 10^3\ s^{-1}$, $k_{rev}[P] = 6.9\ s^{-1}$; Blue: $k_{on} = 10^8\ M^{-1}s^{-1}$, $k_{off} = 10\ s^{-1}$, $k_{cat} = 30\ s^{-1}$, $k_{rev}[P] = 6.2\times10^{-1}\ s^{-1}$; Purple: $k_{on} = 10^8\ M^{-1}s^{-1}$, $k_{off} = 50\ s^{-1}$, $k_{cat} = 10\ s^{-1}$, $k_{rev}[P] = 4.1\times10^{-2}\ s^{-1}$.

where $-\beta\Delta\mu^{l,\mathrm{pol}}_{\mathrm{Gly,Gly1}} = \ln\left(\frac{k_{on}[C]k_{rec,f}k^C_{hyd,f}k^C_{pol}}{k_{off}k^C_{rec,r}k^C_{hyd,r}k^C_{dep}}\right)$, and [C] represents the concentration of the ternary complex (Gly-tRNA$_1^{Gly}$)-(EF-Tu)-GTP (Fig. 3B) The term $\ln(\eta^l_{Gly})$ is required to account for the fact that Gly can be depolymerized at position l only η^l_{Gly} fraction of the time. Similarly, the affinity of incorporating Gly along the near-cognate kinetic path associated with Gly-tRNA$_3^{Gly}$ is

$$\beta\mathcal{A}^{l,\mathrm{pol}}_{\mathrm{Gly,Gly3}} = -\beta\Delta\mu^{l,\mathrm{pol}}_{\mathrm{Gly,Gly3}} - \ln(\eta^l_{\mathrm{Gly}}), \tag{S20}$$

where $-\beta\Delta\mu^{l,\mathrm{pol}}_{\mathrm{Gly,Gly3}} = \ln\left(\frac{k_{on}[NC]k_{rec,f}k^{NC}_{hyd,f}k^{NC}_{pol}}{k_{off}k^{NC}_{rec,r}k^{NC}_{hyd,r}k^{NC}_{dep}}\right)$, and [NC] = [(Gly-tRNA$_3^{Gly}$)-(EF-Tu)-GTP]. Generally, we denote the polymerization affinity of amino acid a associated with aa-tRNA t by $\mathcal{A}^{l,\mathrm{pol}}_{a,t}$.

Next, we let $\langle J^{l,\mathrm{fut}}_{a,t}\rangle$ and $\Delta\mu^{l,\mathrm{fut}}_{a,t}$ be the current and affinity of the futile cycle within the incorporation path of aa-tRNA t associated with amino acid a. Denoting the net polymerization flux by $\langle J^{l,\mathrm{pol}}\rangle \left(= \sum_{a\in\{aa\}}\sum_{t\in T^l_a}\langle J^{l,\mathrm{pol}}_{a,t}\rangle\right)$, we can write the affinity of mRNA translation at position l as

$$\beta\mathcal{A}^l = \underbrace{-\beta\sum_{a\in\{aa\}}\sum_{t\in T^l_a}\left(\frac{\langle J^{l,\mathrm{pol}}_{a,t}\rangle}{\langle J^{l,\mathrm{pol}}\rangle}\Delta\mu^{l,\mathrm{pol}}_{a,t} + \frac{\langle J^{l,\mathrm{fut}}_{a,t}\rangle}{\langle J^{l,\mathrm{pol}}\rangle}\Delta\mu^{l,\mathrm{fut}}_{a,t}\right)}_{=-\beta\Delta\mu} \underbrace{-\sum_{a\in\{aa\}}\eta^l_a\ln\eta^l_a}_{=I}. \tag{S21}$$

For the Bennett model, which involves two types of monomers, with one incorporation pathway each, without any futile cycles, we recover eq. S10. To estimate I in the extended model of translation (Fig. 4), we sum the Shannon-entropy of each position l along the protein sequence.

COMPUTATION OF η AND Q

We will work through the process of calculating η and Q in the T7 DNA polymerase model, for which we apply Koza's method of calculating currents and fluctuations in kinetic networks [9]. To simplify the notation, we will relabel the states in Fig. 2A by indices 1 through 5; i.e., $E^{(1)} \to 1$, $c^{(2)} \to 2$, $c^{(3)} \to 3$, $i^{(2)} \to 4$, and $i^{(3)} \to 5$. Additionally, we will relabel the rate constants so that $k_{\mu,\nu}$ represents the rate of the reaction from state μ to ν; i.e., $k_{1,2} = k^c_{on}$[dNTP], $k_{2,3} = k^c_{conf}$, and so forth. Here, the depolymerization rate constants are set to $k_{1,3} = (1-\eta)k^c_{dep}$ and $k_{1,5} = \eta k^i_{dep}$.

To begin, we will compute the current of correct nucleotide incorporation $\langle J^c_{pol}\rangle$. For the i-th chemical state ($i \in \{1,2,\ldots,N\}$, where $N = 5$) at time t, let $\mu(t) \equiv i + N\times n^c(t)$ be the generalized state of the system after completing $n^c(t)$ correct nucleotide incorporation cycles. Then, let $P(\mu,t)$ be the probability of the system to be in state μ at time t. The time evolution of $P(\mu,t)$ is given by

$$\frac{\partial P(\mu,t)}{\partial t} = \sum_\xi\left[k_{\mu-\xi,\mu}P(\mu-\xi,t) - k_{\mu,\mu-\xi}P(\mu,t)\right], \tag{S22}$$

where the index ξ runs through all states one reaction away from state μ. Here, the periodicity of the network model constrains the rate constants so that $k_{\mu,\nu} = k_{i,j}$ for $\mu = i \pmod N$ and $\nu = j \pmod N$. Following ref. [9], we define $P_j(\mu,t)$ as

$$P_j(\mu,t) \equiv P(\mu,t)\delta^N_{\mu,j}, \tag{S23}$$

where

$$\delta^N_{\mu,j} = \begin{cases} 1, & \text{if } j = \mu \pmod N \\ 0, & \text{otherwise,}\end{cases}$$

for $j \in \{1,2,\ldots,N\}$. By multiplying $\delta^N_{\mu,j}$ to both sides of

eq. S22 and using the equality $\delta^N_{\mu,j} = \delta^N_{\mu-\xi,j-\xi}$, we get

$$\frac{\partial P_j(\mu,t)}{\partial t} = \sum_{\xi}\left[k_{j-\xi,j}P_{j-\xi}(\mu-\xi,t) - k_{j,j-\xi}P_j(\mu,t)\right]. \quad \text{(S24)}$$

To derive the expression of $\langle J^c_{\text{pol}}\rangle$ as a function of the rate constants, we define the generating function

$$\mathcal{G}^c_j(z,t) \equiv \sum_{\mu=-\infty}^{\infty} e^{zX_\mu}P_j(\mu,t), \quad \text{(S25)}$$

where X_μ is the coordinate for the correct nucleotide incorporation cycle at state μ. Then, the time derivative of the generating function can be written as

$$\frac{\partial \mathcal{G}^c_j(z,t)}{\partial t} = \sum_{i=1}^{N}\Gamma^c_{i,j}(z)\mathcal{G}^c_i(z,t), \quad \text{(S26)}$$

where the matrix $\Gamma^c(z)$ is defined as

$$\Gamma^c_{i,j}(z) = \begin{cases} k_{i,j}e^{zd_{i,j}}, & \text{if } i \neq j \\ -\sum_{m=1(\neq i)}^{N} k_{i,m}, & \text{if } i = j, \end{cases}$$

and also shown in the matrix form below

$$\Gamma^c(z) = \begin{bmatrix} -\left(k^c_{\text{on}}[\text{dNTP}] + k^i_{\text{on}}[\text{dNTP}] + k^c_{\text{dep}}(1-\eta) + k^i_{\text{dep}}\eta\right) & k^c_{\text{on}}[\text{dNTP}] & k^c_{\text{dep}}(1-\eta)e^{-z} & k^i_{\text{on}}[\text{dNTP}] & k^i_{\text{dep}}\eta \\ k^c_{\text{off}} & -\left(k^c_{\text{off}} + k^c_{\text{conf},f}\right) & k^c_{\text{conf},f} & 0 & 0 \\ k^c_{\text{pol}}e^{z} & k^c_{\text{conf},r} & -\left(k^c_{\text{pol}} + k^c_{\text{conf},r}\right) & 0 & 0 \\ k^i_{\text{off}} & 0 & 0 & -\left(k^i_{\text{off}} + k^i_{\text{conf},f}\right) & k^i_{\text{conf},f} \\ k^i_{\text{pol}} & 0 & 0 & k^i_{\text{conf},r} & -\left(k^i_{\text{pol}} + k^i_{\text{conf},r}\right) \end{bmatrix} \quad \text{(S27)}$$

For the computation of $\langle J^c_{\text{pol}}\rangle$, $d_{i,j}$ is defined as

$$d_{i,j} = \begin{cases} 1, & \text{if } i = 3 \text{ and } j = 1 \\ -1, & \text{if } i = 1 \text{ and } j = 3 \\ 0, & \text{otherwise.} \end{cases} \quad \text{(S28)}$$

Next, we define $\mathcal{G}^c(z,t) \equiv \sum_{i=1}^{N}\mathcal{G}^c_i(z,t)$ and denote the coordinate of the correct incorporation cycle by $X^c(t)$. Then, it can be shown that

$$\langle J^c_{\text{pol}}\rangle = \lim_{t\to\infty}\frac{\langle X^c(t)\rangle}{t} = \lim_{t\to\infty}\frac{\partial_z\mathcal{G}^c(z,t)|_{z=0}}{t} = \partial_z\Lambda^c_0(z)|_{z=0}, \quad \text{(S29)}$$

and

$$\begin{aligned}\left\langle\left(\delta J^c_{\text{pol}}\right)^2\right\rangle &= \lim_{t\to\infty}\frac{\langle (X^c(t))^2\rangle - \langle X^c(t)\rangle^2}{t} \\ &= \lim_{t\to\infty}\frac{\partial_z^2\mathcal{G}^c(z,t)|_{z=0} - (\partial_z\mathcal{G}^c(z,t)|_{z=0})^2}{t} \\ &= \partial_z^2\Lambda^c_0(z)|_{z=0}, \end{aligned} \quad \text{(S30)}$$

where $\Lambda^c_0(z)$ denotes the maximum eigenvalue of the matrix $\Gamma^c(z)$. Now, let $C_n(z)$ denote the coefficients of the characteristic polynomial of $\Gamma^c(z)$ (i.e., $\sum_{n=0}^{N} C_n(z)\Lambda^c_0(z)^n = 0$). Then, we can write the following expressions for $\langle J^c_{\text{pol}}\rangle$ and $\langle(\delta J^c_{\text{pol}})^2\rangle$,

$$\langle J^c_{\text{pol}}\rangle = \left(\Lambda^c_0\right)' = -\frac{C_0'}{C_1}, \quad \text{(S31)}$$

$$\langle(\delta J^c_{\text{pol}})^2\rangle = \left(\Lambda^c_0\right)'' = -\frac{C_0'' + 2C_1'(\Lambda^c_0)' + 2C_2\left((\Lambda^c_0)'\right)^2}{C_1}. \quad \text{(S32)}$$

where $'$ denotes the derivative with respect to z evaluated at $z = 0$, and C_1 and C_2 are evaluated at $z = 0$. We can analogously compute the current of incorrect monomer incorporation, $\langle J^i_{\text{pol}}\rangle$, by constructing the corresponding matrix $\Gamma^i(z)$, in which the non-diagonal entries of $k_{5,1}$ and $k_{1,5}$ are multiplied by e^z and e^{-z}, respectively. Since $k_{1,3}$ and $k_{1,5}$ are functions of η, $\langle J^c_{\text{pol}}\rangle$ and $\langle J^i_{\text{pol}}\rangle$ are functions of η. Thus, we can solve for η by the equality

$$\eta = \frac{\langle J^i_{\text{pol}}\rangle}{\langle J^i_{\text{pol}}\rangle + \langle J^c_{\text{pol}}\rangle}. \quad \text{(S33)}$$

After we obtain the numerical expression of η, we can construct the matrix $\Gamma(z)$ to calculate the total flux $\langle J_{\text{pol}}\rangle = \langle J^c_{\text{pol}}\rangle + \langle J^i_{\text{pol}}\rangle$ and its fluctuation $\langle(\delta J_{\text{pol}})^2\rangle$. With known values of η and $\langle J_{\text{pol}}\rangle$, the affinity of replication ($\mathcal{A}$) can be computed by eq. S21. In sum, we have demonstrated how to calculate $Q = \mathcal{A}\langle(\delta J_{\text{pol}})^2\rangle/\langle J_{\text{pol}}\rangle$ for the T7 DNA polymerase model.

We can calculate η and Q of the simplified ribosome model in a similar way. To obtain $\langle J^{\text{C}}_{\text{pol}}\rangle$ as a function of η, we construct the corresponding matrix $\Gamma^{\text{C}}(z)$, in which we multiply

the non-diagonal entries corresponding to $k^{\rm C}_{{\rm pol},f}$ and $k^{\rm C}_{{\rm pol},r}$ by e^{z} and e^{-z}, respectively. The expression for $\langle J^{\rm NC}_{\rm pol}\rangle$ is obtained analogously, by constructing the corresponding matrix $\Gamma^{\rm NC}(z)$. After solving for the numerical value of the error probability, we can calculate the total polymerization rate $\langle J_{\rm pol}\rangle$ and its fluctuation $\langle(\delta J_{\rm pol})^2\rangle$ by constructing the corresponding matrix $\Gamma(z)$. When computing the affinity, we must also include the contribution from the futile cycle fluxes. To calculate the futile cycle current $\langle J_{\rm fut}\rangle$, we construct the corresponding matrix $\Gamma^{\rm fut}(z)$, in which we multiply the non-diagonal entries corresponding to $k^{\rm C}_{{\rm PR},f}$, $k^{\rm NC}_{{\rm PR},f}$, $k^{\rm C}_{{\rm PR},r}$ and $k^{\rm NC}_{{\rm PR},r}$ by e^{z}, e^{z}, e^{-z} and e^{-z}, respectively. Finally, the affinity of translation can be computed by eq. S21.

STOCHASTIC SIMULATION OF DNA REPLICATION

We simulated the replication of the first 300 base pairs of the T7 DNA polymerase gene sequence at the single molecule level, with Gillespie's algorithm [10]. The simulation starts with the polymerase in the apo state at the beginning of the gene sequence. The only reactions available at this state are the binding reactions of the 4 dNTPs, which are assumed to be at equal concentrations. After the binding of a dNTP, the simulation trajectories are generated based on the kinetic network shown in Fig. 2A. After each polymerization reaction, the polymerase translocates on the DNA and reads the next nucleotide. The simulation is terminated when the 300$^{\rm th}$ nucleotide of the gene sequence is polymerized.

The dynamics of DNA replication simulations are studied using the ensemble of trajectories generated. The total number of steps (N_{rep}) in completing the replication of the DNA sequence varies from one realization to another. Selecting the completion time of replication ($\mathcal{T}$) as the output observable for each realization, we define TUR of replication as

$$\mathcal{Q} = \left[-\Delta\mu + \beta^{-1} I\right] \frac{\langle(\delta\mathcal{T})^2\rangle}{\langle\mathcal{T}\rangle^2}, \tag{S34}$$

where the dissipation has contributions from the free energy drive ($\Delta\mu$) and Shannon-entropy (I). Denoting the forward and reverse rate constants of each kinetic step by $k_{i,f}$ and $k_{i,r}$ for $i = 1,\ldots,N_{rep}$, we can compute the average free energy drive by $-\beta\Delta\mu = \langle\sum_{i=1}^{N_{rep}} \ln\frac{k_{i,f}}{k_{i,r}}\rangle$, where $\langle\ldots\rangle$ denotes the average over the ensemble of 10^4 realizations. The entropic contribution is computed as $I = -\sum_{l=1}^{300}\sum_{i_{\rm dNTP}=1}^{4}\eta^{l}_{i_{\rm dNTP}}\ln\eta^{l}_{i_{\rm dNTP}}$, where $\eta^{l}_{i_{\rm dNTP}}$ is the probability of incorporating one of the 4 types of dNTPs, at the l-th position.

MODEL OF TRANSLATION

We provide more details on the model of translation by the ribosome in Fig. 3. Our model is a modified version of that from Rudorf *et. al.* [11], in which we combine all linear chains of consecutive and irreversible reactions into single reactions. For instance, consider two consecutive and irreversible reactions (1) $\rightarrow$ (2) and (2) $\rightarrow$ (3), with respective rate constants k_{12} and k_{23}, defined among three states (1), (2) and (3). If there are no other reactions associated with the state (2), we remove the state (2), and define a new reaction (1) $\rightarrow$ (3) with the rate constant $k_{13}^{-1} \equiv k_{12}^{-1} + k_{23}^{-1}$.

To calculate the entropy productions, we defined reverse rate constants for all reactions, which were constrained by the affinity of the corresponding kinetic cycle. We assumed that affinities of the parallel kinetic cycles for the cognate and near-cognate aa-tRNAs were identical. The affinity of the futile cycle ($\Delta\mu_{\rm fut}$) comes from GTP hydrolysis, which dissipates $\approx$ 20 k_BT [12]. Therefore, we applied the following constraints, where $-\beta\Delta\mu_{\rm fut} = 20$.

$$\ln\left(\frac{k_{\rm on}[{\rm C}]k_{{\rm rec},f}k^{\rm C}_{{\rm hyd},f}k^{\rm C}_{{\rm PR},f}}{k_{\rm off}k^{\rm C}_{{\rm rec},r}k^{\rm C}_{{\rm hyd},r}k^{\rm C}_{{\rm PR},r}[{\rm C}']}\right) = -\beta\Delta\mu_{\rm fut}, \tag{S35}$$

$$\ln\left(\frac{k_{\rm on}[{\rm NC}]k_{{\rm rec},f}k^{\rm NC}_{{\rm hyd},f}k^{\rm NC}_{{\rm PR},f}}{k_{\rm off}k^{\rm NC}_{{\rm rec},r}k^{\rm NC}_{{\rm hyd},r}k^{\rm NC}_{{\rm PR},r}[{\rm NC}']}\right) = -\beta\Delta\mu_{\rm fut}. \tag{S36}$$

The affinity involved with polymerization ($\Delta\mu_{\rm pol}$) can be estimated as the sum of the free energies of GTP hydrolysis, peptide bond formation, and the cleavage of the ester bond between the tRNA and the amino acid. The hydrolysis of the GTP molecule incurs the dissipation of $\approx 20\,k_BT$. Conversely, each peptide bond synthesized stores $\approx 5\,k_BT$ of free energy [13]. The standard free energy of the ester bond between the amino acid and the tRNA is $\approx 12\,k_BT$ [12, 14]. Since the ratio of charged to uncharged tRNAs is ≈ 10 fold [15], the net free energy of the ester bond between the amino acid and tRNA is $\approx 15\,k_BT$. Then, $-\beta\Delta\mu_{\rm pol} \approx 30$, which gives the following constraints

$$\ln\left(\frac{k_{\rm on}[{\rm C}]k_{{\rm rec},f}k^{\rm C}_{{\rm hyd},f}k^{\rm C}_{\rm pol}}{k_{\rm off}k^{\rm C}_{{\rm rec},r}k^{\rm C}_{{\rm hyd},r}k^{\rm C}_{\rm dep}}\right) = -\beta\Delta\mu_{\rm pol}, \tag{S37}$$

$$\ln\left(\frac{k_{\rm on}[{\rm NC}]k_{{\rm rec},f}k^{\rm NC}_{{\rm hyd},f}k^{\rm NC}_{\rm pol}}{k_{\rm off}k^{\rm NC}_{{\rm rec},r}k^{\rm NC}_{{\rm hyd},r}k^{\rm NC}_{\rm dep}}\right) = -\beta\Delta\mu_{\rm pol}. \tag{S38}$$

The terms $k^{\rm C}_{\rm dep}$ and $k^{\rm NC}_{\rm dep}$ implicitly take into account the concentration of tRNA and (EF-Tu)-GDP, which detach during the final polymerization step. In order to fully constrain all the rate constants, we set $k^{\rm C}_{{\rm hyd},r} = 10^{-3}k^{\rm C}_{{\rm hyd},f}$ and $k^{\rm NC}_{{\rm hyd},r} = 10^{-3}k^{\rm NC}_{{\rm hyd},f}$. Modest changes to these affinity related constraints (eqs S35-S38) do not affect the qualitative conclusions of our work.

CALCULATION OF THE TERNARY COMPLEX CONCENTRATION

The ternary complex concentration was modeled as a function of the concentration of its components, aa-tRNA, EF-Tu, GTP, and GDP. First, EF-Tu binds with GTP and GDP to

form (EF-Tu)-GTP and (EF-Tu)-GDP, respectively. Then, aa-tRNA binds with (EF-Tu)-GTP and (EF-Tu)-GDP to form (aa-tRNA)-(EF-Tu)-GTP and (aa-tRNA)-(EF-Tu)-GDP, respectively [16]. Here, (aa-tRNA)-(EF-Tu)-GTP and (aa-tRNA)-(EF-Tu)-GDP represent the combined total of all the respective 42 individual ternary complexes. Assuming equilibrium, we can write the following equalities,

$$[\text{(aa-tRNA)-(EF-Tu)-GTP}] = \frac{[\text{aa-tRNA}][\text{(EF-Tu)-GTP}]}{K_{\text{aaGTP}}}, \tag{S39}$$

$$[\text{(aa-tRNA)-(EF-Tu)-GDP}] = \frac{[\text{aa-tRNA}][\text{(EF-Tu)-GDP}]}{K_{\text{aaGDP}}}, \tag{S40}$$

$$[\text{(EF-Tu)-(GTP)}] = \frac{[\text{EF-Tu}][\text{GTP}]}{K_{\text{GTP}}}, \tag{S41}$$

$$[\text{(EF-Tu)-(GDP)}] = \frac{[\text{EF-Tu}][\text{GDP}]}{K_{\text{GDP}}}, \tag{S42}$$

where the respective dissociation constants are set to $K_{\text{aaGTP}} = 10^{-1}\,\mu$M [16], $K_{\text{aaGDP}} = 14\,\mu$M [16], $K_{\text{GTP}} = 6\times10^{-2}\ \mu M$ [17], and $K_{\text{GDP}} = 10^{-3}\,\mu$M [17]. Using eqs. S39-S42, we can solve for the concentration of all chemical species given the total [EF-Tu], [aa-tRNA], [GTP], and [GDP]. Unless specified otherwise, all the ribosome model plots (Fig. 3 and Fig. 4) are made at the cellular condition, with [EF-Tu] = 250 μM [18], [aa-tRNA] $\approx$ 200 μM [19], [GTP] = 5000 μM [20], and [GDP] = 700 μM [20]. At the cellular condition, the concentration of the GTP bound ternary complex is [(aa-tRNA)-(EF-Tu)-GTP] $\approx$ [aa-tRNA] $\approx$ 200 μM.

Assuming that all species of aa-tRNA were bound to (EF-Tu)-GTP and (EF-Tu)-GDP with equal binding constants K_{aaGTP} and K_{aaGDP}, we computed the concentration of individual ternary complexes by referencing the measured concentration ratios among individual aa-tRNA species. For instance, the concentration of the cognate ternary complexes of codon UUU, which encodes for Phe, are

$$[\text{C}] = \frac{[\text{Phe-tRNA}^{\text{Phe}}]_{\text{WT}}}{[\text{aa-tRNA}]_{\text{WT}}}[\text{(aa-tRNA)-(EF-Tu)-GTP}], \tag{S43}$$

$$[\text{C}'] = \frac{[\text{Phe-tRNA}^{\text{Phe}}]_{\text{WT}}}{[\text{aa-tRNA}]_{\text{WT}}}[\text{(aa-tRNA)-(EF-Tu)-GDP}], \tag{S44}$$

where the subscript WT represents the cellular concentrations obtained from ref. [19].

EXTENDED MODEL OF TRANSLATION

Excluding the three stop codons, there are 61 types of codons encoding for 20 amino acids. For *E. coli*, the 43 types of tRNAs were identified by Dong *et. al* [19]. Out of these, the pairs Gly1-Gly2 and Ile1-Ile2 were not differentiated in the concentration measurements. In our simulations, we assumed that Gly1 and Gly2 (resp. Ile1 and Ile2) were each present in the cellular milieu at half of the measured concentration of the Gly1-Gly2 pair (resp. Ile1-Ile2 pair). We removed the seleno-cysteine carrying tRNA from the analysis, since it is low in concentration, and it is incorporated into the polypeptide through a different kinetic scheme from the rest of the aa-tRNAs. Overall, we included total 42 types of aa-tRNAs in the extended translation model, with measurements from *E. coli* dividing every $\approx$ 86 minutes [19]. The cognate, near-cognate, and non-cognate groupings of aa-tRNAs for each codon is shown in Fig. S7.

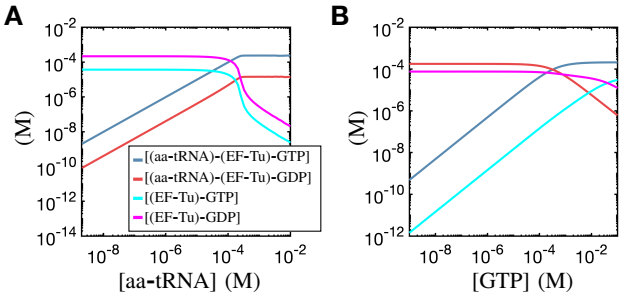

FIG. S3. The concentration of the components of the ternary complex as functions of [aa-tRNA] and [GTP]. (A) The dependence of [(aa-tRNA)-(EF-Tu)-GTP], [(aa-tRNA)-(EF-Tu)-GDP], [(EF-Tu)-GTP], and [(EF-Tu)-GDP] on [aa-tRNA], where [GTP] was fixed at 5 mM. (B) The dependence of [(aa-tRNA)-(EF-Tu)-GTP], [(aa-tRNA)-(EF-Tu)-GDP], [(EF-Tu)-GTP], and [(EF-Tu)-GDP] on [GTP], where [aa-tRNA] was fixed at $\approx$ 0.2mM. The curves are color-coded identically as those in (A). For both plots, we assumed cellular levels of [GDP](=0.7 mM) and [EF-TU](=0.25 mM), and used equilibrium dissociation constants as described in the SI text.

mRNA translation by the ribosome at the single molecule level is simulated with Gillespie's algorithm [10]. The simulation starts with the ribosome in the apo state at the start codon. The only reactions available at this state are the bindings of the 42 aa-tRNAs, the concentrations of which were taken from Dong *et. al.* [19]. After the binding of an aa-tRNA, the simulation trajectories were generated on the kinetic network shown in Fig. 3B. After each polymerization reaction, the ribosome reads the next codon, translocating along the mRNA. The simulation is terminated when the ribosome completes the translation of the last codon.

HOPFIELD MODEL

We provide more details on the modified Hopfield model [21]. The reaction cycle of the Hopfield model is composed of substrate binding ($\mathbf{E}+\mathbf{C} \rightleftharpoons \mathbf{EC}$ and $\mathbf{E}+\mathbf{I} \rightleftharpoons \mathbf{EI}$), followed by the effectively irreversible steps of ATP hydrolysis ($\mathbf{EC} \rightleftharpoons \mathbf{EC}^*$ and $\mathbf{EI} \rightleftharpoons \mathbf{EI}^*$) and polymerization ($\mathbf{EC}^* \rightleftharpoons \mathbf{E}$ and $\mathbf{EI}^* \rightleftharpoons \mathbf{E}$) (Fig. S4). At states $\mathbf{EC}^*$ and $\mathbf{EI}^*$, the substrate can also dissociate through the proofreading reaction (PR, $\mathbf{EC}^* \rightleftharpoons \mathbf{E}+\mathbf{C}$ and $\mathbf{EI}^* \rightleftharpoons \mathbf{E}+\mathbf{I}$), which is also effec-

tively irreversible.

In the original formulation of the Hopfield model, the forward kinetic rate constants of correct and incorrect pathways were identical. In our modified version, the forward constants satisfy the following relations

$$e^{\beta\delta} = \frac{k^c_{\rm on}}{k^i_{\rm on}} = \frac{k^c_{{\rm hyd},f}}{k^i_{{\rm hyd},f}} = \frac{k^c_{\rm pol}}{k^i_{\rm pol}} = \frac{k^c_{{\rm PR},r}}{k^i_{{\rm PR},r}}, \tag{S45}$$

with $\beta\delta > 0$. Next, to allow for error reduction by proofreading, we set the forward kinetic rates so that $k^c_{\rm pol} \ll k^c_{{\rm hyd},f} \ll k^c_{\rm on}$. Finally, we constrained the reverse reaction rates so that the affinities associated with the kinetic cycles are $\Delta\mu^c_{\rm pol} = -20\ k_BT$, and $\Delta\mu^i_{\rm pol} = \Delta\mu^c_{\rm fut} = \Delta\mu^i_{\rm fut} = -18\ k_BT$. The rate constants used to generate Fig. 5 and Fig. S9 are provided in Table S3.

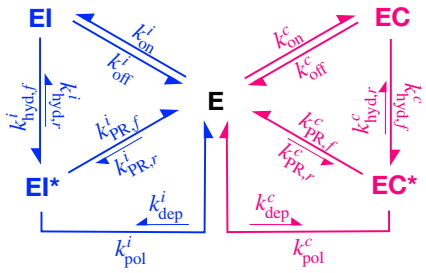

FIG. S4. Schematic of the modified Hopfield model [21].

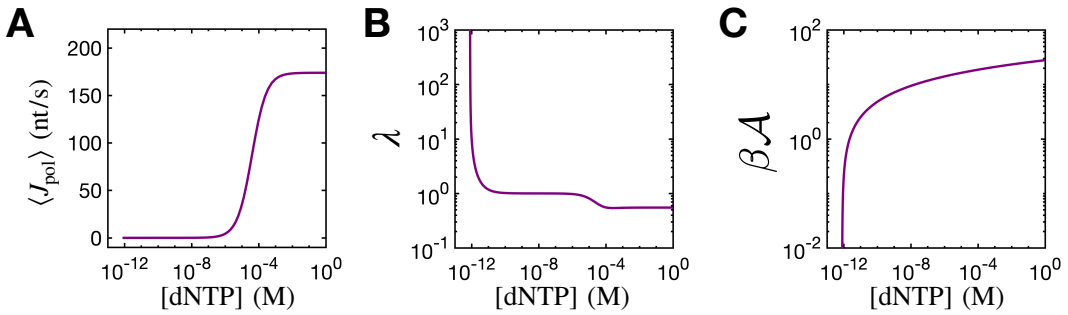

FIG. S5. Dynamical properties of the exonuclease-deficient T7 DNA polymerase, obtained at identical conditions as in Fig. 2. (A) The reaction current ($\langle J_{\rm pol}\rangle$), (B) Fano factor of the reaction current (λ), and (C) the affinity ($\mathcal{A}$) are plotted against [dNTP].

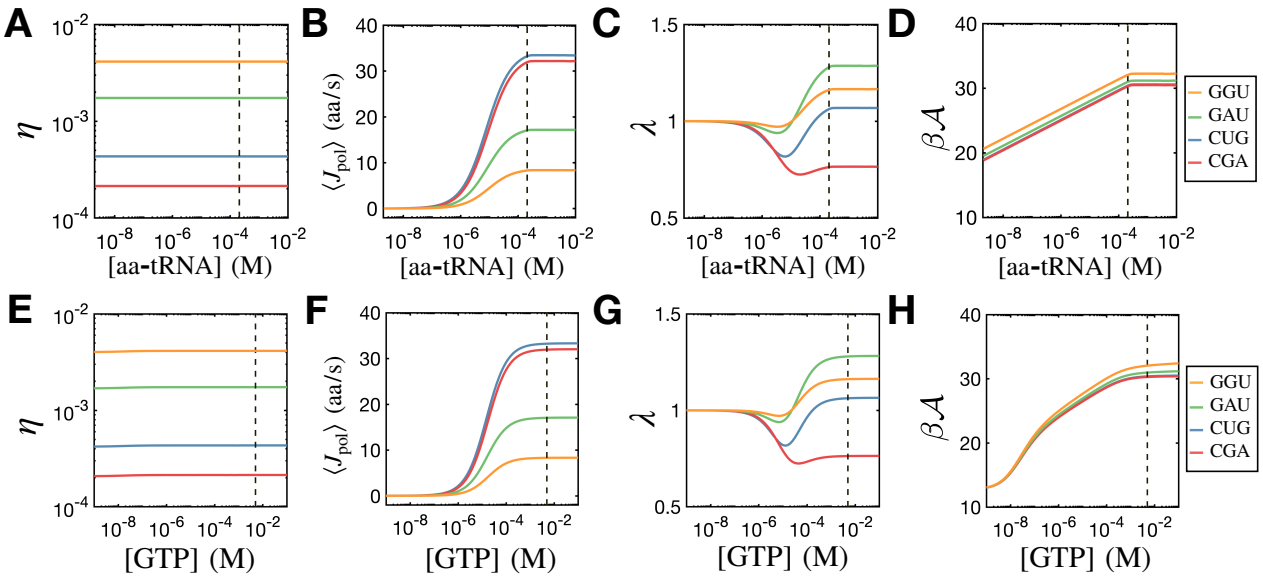

FIG. S6. Dynamical properties of mRNA translation by *E. coli* ribosome. (A) The error probability (η), (B) polymerization current ($\langle J_{\rm pol}\rangle$), (C) Fano factor of the polymerization current (λ), and (D) affinity ($\mathcal{A}$) plotted against [aa-tRNA], at identical conditions as in Fig. 3D. (E) The error probability (η), (F) polymerization current ($\langle J_{\rm pol}\rangle$), (G) Fano factor of the polymerization current (λ), and (H) affinity ($\mathcal{A}$) plotted against [GTP], at identical conditions as in Fig. 3E. For all plots, the dashed black line represents the cellular condition in *E. coli*.

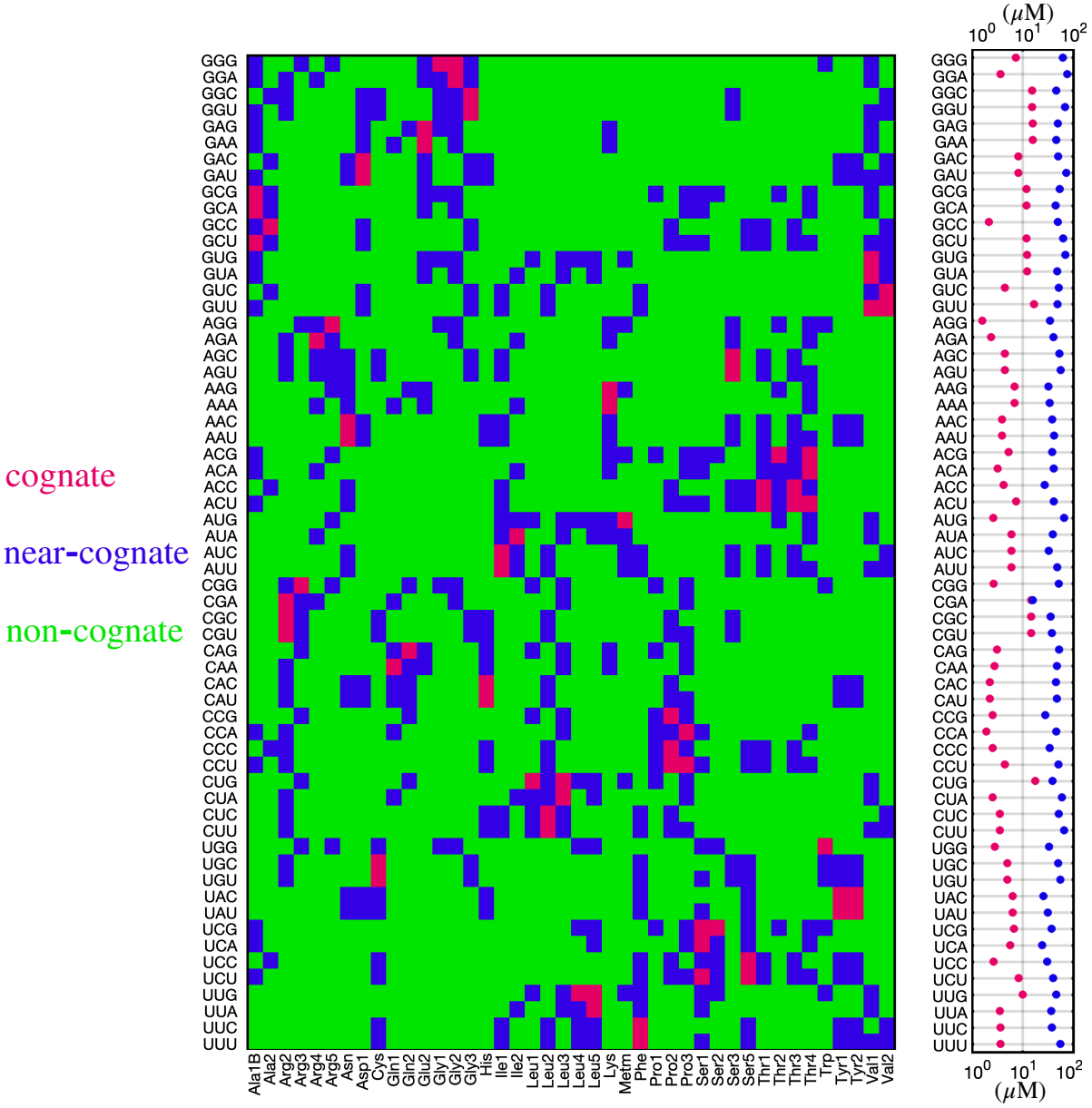

FIG. S7. (Left) The groupings of cognate (red), near-cognate (blue), and non-cognate (green) aa-tRNAs for each codon [11]. (Right) The sum of the concentration of the cognate (red) and near-cognate (blue) aa-tRNA species, plotted for each codon [19].

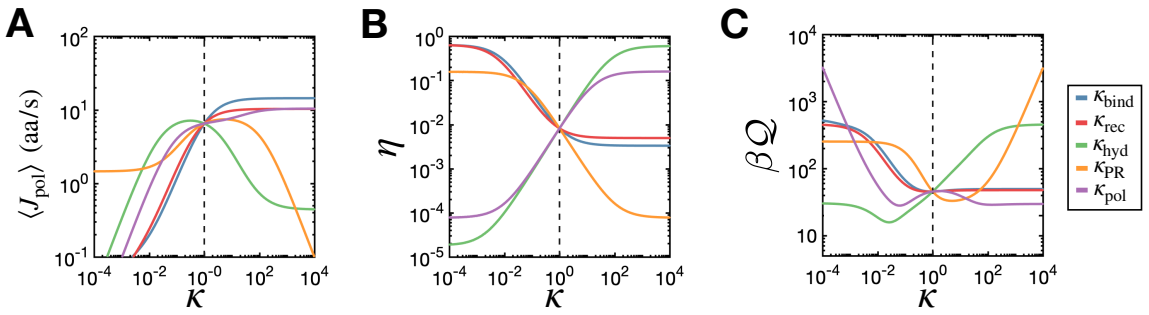

FIG. S8. Dynamical properties of mRNA translation by the *E. coli* ribosome with respect to perturbations to the wild type rate constants. (A) The polymerization current ($\langle J_{\rm pol}\rangle$), (B) the error probability (η), and (C) Q as functions of $\kappa_{\rm bind}$, $\kappa_{\rm rec}$, $\kappa_{\rm hyd}$, $\kappa_{\rm PR}$, and $\kappa_{\rm pol}$. The perturbative parameters $\kappa_{\rm bind}$, $\kappa_{\rm rec}$, $\kappa_{\rm hyd}$, $\kappa_{\rm PR}$, and $\kappa_{\rm pol}$ were each multiplied to the reactions associated with binding, codon-recognition, GTP-hydrolysis, proofreading, and polymerization, respectively. The black dashed lines represent the wild type condition for the codon UUU.

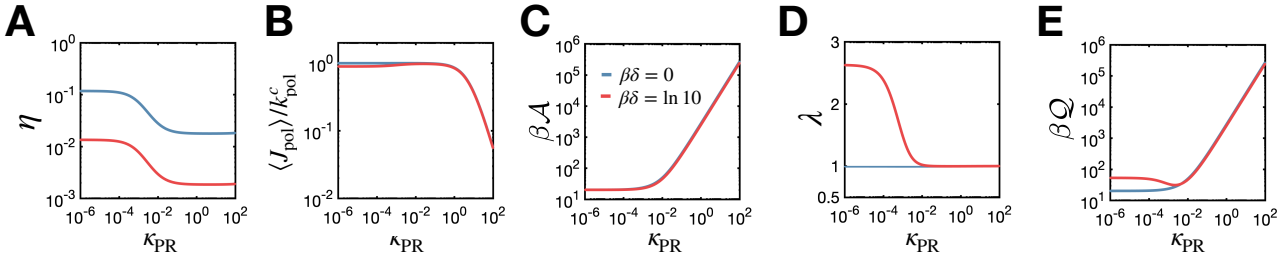

FIG. S9. Dynamical properties of the modified Hopfield model with kinetic discrimination. (A) The error probability (η), (B) normalized polymerization current ($\langle J_{\rm pol}\rangle/k^c_{\rm pol}$), (C) affinity ($\mathcal{A}$), (D) Fano factor ($\lambda$), and (E) Q as functions of $\kappa_{\rm PR}$. For (A)-(E) The perturbative parameter $\kappa_{\rm PR}$ is multiplied to the rates $k^c_{{\rm PR},f}$, $k^i_{{\rm PR},f}$, $k^c_{{\rm PR},r}$ and $k^i_{{\rm PR},r}$. The blue line presents the original Hopfield model with $\beta\delta = 0$, and the orange line represents the modified Hopfield model with kinetic discrimination, with $\beta\delta = \ln 10$. In (C), the orange and blue lines are nearly identical. The rate constants used to make these plots are given in Table S3.

TABLE S1. Rate constants for the T7 DNA polymerase, from ref. [22]. The rate constants $k^c_{\rm dep}$ and $k^i_{\rm dep}$ are determined such that the affinities of correct and incorrect monomer incorporations are 20 and 15 k_BT, respectively, when [dNTP]=100 μM [23, 24]. The terms $k^c_{\rm dep}$ and $k^i_{\rm dep}$ implicitly take into account the concentration of PPi which detaches during the final polymerization step.

$k^c_{\rm on}$	$10^2\ \mu{\rm M}^{-1}s^{-1}$
$k^c_{{\rm conf},f}$	$6.0\times10^2\ s^{-1}$
$k^c_{\rm pol}$	$3.6\times10^2\ s^{-1}$
$k^c_{\rm off}$	$2.8\times10^3\ s^{-1}$
$k^c_{{\rm conf},r}$	$1.6\ s^{-1}$
$k^c_{\rm dep}$	$10^{-3}\ s^{-1}$
$k^i_{\rm on}$	$10^2\ \mu{\rm M}^{-1}s^{-1}$
$k^i_{{\rm conf},f}$	$2.2\times10^2\ s^{-1}$
$k^i_{\rm pol}$	$3.0\times10^{-1}\ s^{-1}$
$k^i_{\rm off}$	$2.0\times10^4\ s^{-1}$
$k^i_{{\rm conf},r}$	$4.2\times10^2\ s^{-1}$
$k^i_{\rm dep}$	$2.4\times10^{-8}\ s^{-1}$

TABLE S2. Rate constants for the wild type *E. coli* ribosome, from ref. [11]. The rate constants $k^{\rm C}_{{\rm PR},r}$, $k^{\rm C}_{\rm dep}$, $k^{\rm NC}_{{\rm PR},r}$, and $k^{\rm NC}_{\rm dep}$ were determined from the constraints associated with the affinity of the kinetic cycles at the wild type condition (eqs S35-S38).

$k_{\rm on}$	$9.4\times10\ \mu{\rm M}^{-1}s^{-1}$
$k_{\rm off}$	$1.4\times10^3\ s^{-1}$
$k_{{\rm rec},f}$	$2.1\times10^3\ s^{-1}$
$k^{\rm C}_{{\rm rec},r}$	$2\ s^{-1}$
$k^{\rm C}_{{\rm hyd},f}$	$3.75\times10^2\ s^{-1}$
$k^{\rm C}_{{\rm hyd},r}$	$k^{\rm C}_{{\rm hyd},f}\times10^{-3}$
$k^{\rm C}_{{\rm PR},f}$	$1\ s^{-1}$
$k^{\rm C}_{\rm pol}$	$1.1\times10^2\ s^{-1}$
$k^{\rm NC}_{{\rm rec},r}$	$2.7\times10^3\ s^{-1}$
$k^{\rm NC}_{{\rm hyd},f}$	$4.9\ s^{-1}$
$k^{\rm NC}_{{\rm hyd},r}$	$k^{\rm NC}_{{\rm hyd},f}\times10^{-3}$
$k^{\rm NC}_{{\rm PR},f}$	$6\ s^{-1}$
$k^{\rm NC}_{\rm pol}$	$2.7\times10^{-1}\ s^{-1}$

TABLE S3. Rate constants for the Hopfield model with kinetic discrimination, where $-\beta\Delta\mu_c = 2$ and $-\beta\Delta\mu_i = 0$. The parameters δ and $\kappa_{\rm PR}$ are as defined in the main text.

$k^c_{\rm on}$	$10^3\ s^{-1}$
$k^c_{\rm off}$	$k^c_{\rm on}e^{\beta\Delta\mu_c}$
$k^c_{{\rm hyd},f}$	$1\ s^{-1}$
$k^c_{{\rm hyd},r}$	$k^c_{{\rm hyd},f}e^{-10}$
$k^c_{\rm pol}$	$10^{-3}\ s^{-1}$
$k^c_{\rm dep}$	$k^c_{\rm pol}e^{-8}$
$k^c_{{\rm PR},f}$	$\kappa_{\rm PR}e^{\beta\Delta\mu_c}\ s^{-1}$
$k^c_{{\rm PR},r}$	$k^c_{{\rm PR},f}e^{-8-\beta\Delta\mu_c}$
$k^i_{\rm on}$	$k^c_{\rm on}e^{-\beta\delta}$
$k^i_{\rm off}$	$k^i_{\rm on}e^{\beta\Delta\mu_i}$
$k^i_{{\rm hyd},f}$	$k^c_{{\rm hyd},f}e^{-\beta\delta}$
$k^i_{{\rm hyd},r}$	$k^i_{{\rm hyd},f}e^{-10}$
$k^i_{\rm pol}$	$k^c_{\rm pol}e^{-\beta\delta}$
$k^i_{\rm dep}$	$k^i_{\rm pol}e^{-8}$
$k^i_{{\rm PR},f}$	$\kappa_{\rm PR}e^{-\beta\delta+\beta\Delta\mu_i}\ s^{-1}$
$k^i_{{\rm PR},r}$	$k^i_{{\rm PR},f}e^{-8-\beta\Delta\mu_i}\ s^{-1}$

[1] C. H. Bennett, BioSystems **11**, 85 (1979).
[2] F. Cady and H. Qian, Phys. Biol. **6** (2009), 10.1088/1478-3975/6/3/036011.
[3] P. Gaspard and D. Andrieux, J. Chem. Phys. **141** (2014), 10.1063/1.4890821.
[4] P. Gaspard, Phys. Rev. E **93** (2016), 10.1103/PhysRevE.93.042419.
[5] P. Gaspard, arXiv:2001.04923 , 1 (2020), arXiv:2001.04923.
[6] W. Hwang and C. Hyeon, J. Phys. Chem. Lett. **9**, 513 (2018).
[7] A. Dechant and S.-i. Sasa, J. Stat. Mech.: Theory Exp. **2018**, 063209 (2018).
[8] S. Pigolotti and P. Sartori, J. Stat. Phys. **162**, 1167 (2016).
[9] Z. Koza, J. Phys. A. **32**, 7637 (1999).
[10] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[11] S. Rudorf, M. Thommen, M. V. Rodnina, and R. Lipowsky, PLoS Comput. Biol. **10** (2014), 10.1371/journal.pcbi.1003909.
[12] J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed. (W. H. Freeman and Company, 2002).
[13] B. R. Martin, Biopolymers **45**, 351 (1998).
[14] R. B. Loftfield, Prog. Nucleic Acid Res. Mol. Biol. **12**, 87 (1972).
[15] M. E. Evans, W. C. Clark, G. Zheng, and T. Pan, Nucleic Acids Res. **45**, e133 (2017).
[16] G. Romero, V. Chau, and R. L. Biltonen, J. Biol. Chem. **260**, 6167 (1985).
[17] K. B. Gromadski, H. J. Wieden, and M. V. Rodnina, Biochemistry **41**, 162 (2002).
[18] Y. Ishihama, T. Schmidt, J. Rappsilber, M. Mann, F. U. Harlt, M. J. Kerner, and D. Frishman, BMC Genomics **9**, 1 (2008).
[19] H. Dong, L. Nilsson, and C. G. Kurland, J. Mol. Biol. **260**, 649 (1996).
[20] B. D. Bennett, E. H. Kimball, M. Gao, R. Osterhout, S. J. Van Dien, and J. D. Rabinowitz, Nat. Chem. Biol. **5**, 593 (2009).
[21] J. Hopfield, Proc. Natl. Acad. Sci. **71**, 4135 (1974).
[22] Y. C. Tsai and K. A. Johnson, Biochemistry **45**, 9675 (2006).
[23] M. F. Goodman, Proc. Natl. Acad. Sci. U. S. A. **94**, 10493 (1997).
[24] C. A. Minetti, D. P. Remeta, H. Miller, C. A. Gelfand, G. E. Plum, A. P. Grollman, and K. J. Breslauer, Proc. Natl. Acad. Sci. U. S. A. **100**, 14719 (2003).