Elucidating interplay of speed and accuracy in biological error correction

Kinshuk Banerjee[a], Anatoly B. Kolomeisky[a,b,1,2], and Oleg A. Igoshin[a,c,1,2]

[a]Center for Theoretical Biological Physics, Rice University, Houston, TX 77005; [b]Department of Chemistry, Rice University, Houston, TX 77005; and [c]Department of Bioengineering, Rice University, Houston, TX 77005

Edited by William Bialek, Princeton University, Princeton, NJ, and approved April 7, 2017 (received for review September 4, 2016)

One of the most fascinating features of biological systems is the ability to sustain high accuracy of all major cellular processes despite the stochastic nature of underlying chemical processes. It is widely believed that such low error values are the result of the error-correcting mechanism known as kinetic proofreading. However, it is usually argued that enhancing the accuracy should result in slowing down the process, leading to the so-called speed–accuracy trade-off. We developed a discrete-state stochastic framework that allowed us to investigate the mechanisms of the proofreading using the method of first-passage processes. With this framework, we simultaneously analyzed the speed and accuracy of the two fundamental biological processes, DNA replication and tRNA selection during the translation. The results indicate that these systems tend to optimize speed rather than accuracy, as long as the error level is tolerable. Interestingly, for these processes, certain kinetic parameters lay in the suboptimal region where their perturbations can improve both speed and accuracy. Additional constraints due to the energetic cost of proofreading also play a role in the error correcting process. Our theoretical findings provide a microscopic picture of how complex biological processes are able to function so fast with high accuracy.

kinetic proofreading mechanisms | stochastic models | first-passage processes

Biological systems exhibit remarkable accuracy in selecting the right substrate from the pool of chemically similar molecules. This property is common to all fundamental biological processes such as DNA replication, RNA transcription, and protein translation (1). The level of fidelity in various stages of genetic information flow depends on their relative importance in sustaining system stability. DNA replication is thought to be the most accurate process, with an error rate $\eta \approx 10^{-8}-10^{-10}$ (2, 3); i.e., only 1 out of 10^8 to 10^{10} incorporated nucleotides is mismatched. RNA transcription ($\eta \approx 10^{-4}-10^{-5}$) and protein translation ($\eta \approx 10^{-3}-10^{-4}$) processes are also quite accurate, but to a somewhat lower degree (4, 5). Failure to maintain such accuracy adversely affects cell viability and survival. For example, mutations affecting the fidelity in translation increase the amount of unfolded proteins leading to apoptosis (6) and to erroneous replication of genetic material (7).

Initially, it was unclear how the small differences in equilibrium binding stability of structurally similar substrates can allow such a high degree of discrimination (8). Then, an explanation was provided independently by Hopfield (9) and Ninio (10), who proposed an error-correction mechanism called kinetic proofreading (KPR). KPR allows enzymes to use the free energy difference between right and wrong substrates multiple times using additional steps (9); this amplifies the small energetic discrimination and results in a lower error compared with that in chemical equilibrium. However, such processes require significant energy consumption (9). To this end, enzymes use some energy-rich molecules, like ATP, to provide for the necessary driving (11, 12). The mechanism was experimentally verified later in different biological systems (13–17). Several recent studies generalized it to more complex networks and found analogies between proofreading and other phenomena such as microtubule growth (18) or bacterial chemotaxis (19). These results broaden the concept of KPR and show that such chemically driven regulatory mechanisms are widely present.

Cells must process genetic information not just accurately but also sufficiently rapidly. Proofreading enhances the accuracy by resetting the system to its initial configuration without progressing to product state (9). The completion time of the reaction is, thus, expected to increase. Hence, there could be a compromise, or trade-off, between accuracy and speed of the process (20). The understanding on this trade-off is mainly based on the Michaelis–Menten (MM) description of specificity (21, 22). These studies indicate that the minimum possible error is achieved at vanishingly low catalytic rate, i.e., when the process is the slowest (9, 21). In contrast, biological polymerization reactions must occur reasonably fast (15, 23). A recent study demonstrated a new speed–accuracy regime in the KPR model by modifying the catalytic rate (18). In this regime, a large gain in speed comes with a relatively small loss in accuracy. The authors suggested that biological systems may use this regime (18). For example, in the tRNA selection process, a fast GTP hydrolysis step speeds up protein synthesis but prevents maximal possible selectivity of the initial tRNA–ribosome binding step (21, 24).

Despite the number of studies, a clear quantitative picture of how the balance between speed and accuracy is tuned is lacking. Several current models of proofreading still mainly focus on the initial stages of substrate selection (22, 25, 26) or assume disparity of rate constants of only a few types of steps (18, 19). In contrast, experimental data show that biological systems have different rates for the right and wrong substrates for each step of the network (4, 14, 15). It is not clear how such distributed

Significance

Biological processes are unique in showing a remarkable level of accuracy in discriminating between similar molecules. This characteristic is attributed to an error-correcting mechanism known as kinetic proofreading. It is widely believed that the enhancement of the accuracy in biological processes always slows them down. By analyzing the fundamental processes of DNA replication and protein translation, we established that these systems maximize speed rather than accuracy with additional energetic constraints. Our theoretical study further indicates that both speed and accuracy can be enhanced in certain parameter regimes. The resulting findings provide a microscopic picture of how complex biological processes can be accomplished so quickly with minimal errors.

Author contributions: A.B.K. and O.A.I. designed research; K.B., A.B.K., and O.A.I. performed research; and K.B., A.B.K., and O.A.I. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]O.A.I. and A.B.K. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: tolya@rice.edu or igoshin@rice.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1614838114/-/DCSupplemental.

discrimination of the reaction rates affects the trade-off. Moreover, proofreading steps come with an extra energy cost to gain higher accuracy (11), but the role of this cost in the trade-off is not apparent. Therefore, to understand the fundamental mechanisms of proofreading in real biological systems, one needs to answer the following questions: (*i*) How does the system set its priorities when choosing between accuracy and speed, two seemingly opposite objectives? (*ii*) Can speed and accuracy change in the same direction; in other words, can perturbations of a kinetic parameter from its naturally selected value improve both speed and accuracy? (*iii*) How does the extra energy expenditure due to KPR affect the speed–accuracy optimization?

Here we focus on the role of reaction kinetics in governing the speed–accuracy trade-off. To this end, we develop a generalized framework to study one-loop KPR networks, assuming distinct rate constants for every step of the right (R) and wrong (W) pathways. Based on this approach, we model the overall selection of the correct substrate over the incorrect one as a first-passage problem, to obtain a full dynamic description of the process (27, 28). This general framework is applied to two important examples, namely, DNA replication by T7 DNA polymerase (DNAP) (3, 14) and protein synthesis by *Escherichia coli* ribosome (22, 29) (Fig. 1 *A* and *B*)). Starting from the experimentally measured rate constants for each system, we vary their values to analyze the resulting changes in speed and accuracy and to assess the trade-off. The role played by the extra energy consumption or cost of proofreading (11) is also investigated. By comparing the behavior of the two systems, we search for general properties of biological error correction.

Methods

Proofreading Networks of Replication and Translation. DNA replication as well as protein synthesis use nucleotide complementarity to select the cognate substrate over other near/noncognate substrates. During replication, dNTP molecules complementary to the DNA template are chosen. Similarly, during protein synthesis, aminoacyl(aa)-tRNAs are picked by ribosome based on the complementarity of their anticodon to the mRNA codon. Wrong substrates that bind initially can be removed by error-correction proofreading mechanisms. Kinetic experiments coupled with modeling revealed a lot of mechanistic details about both of the processes (3, 14, 15, 24). The schemes depicted in Fig. 1 represent the key steps to understand the KPR in these networks.

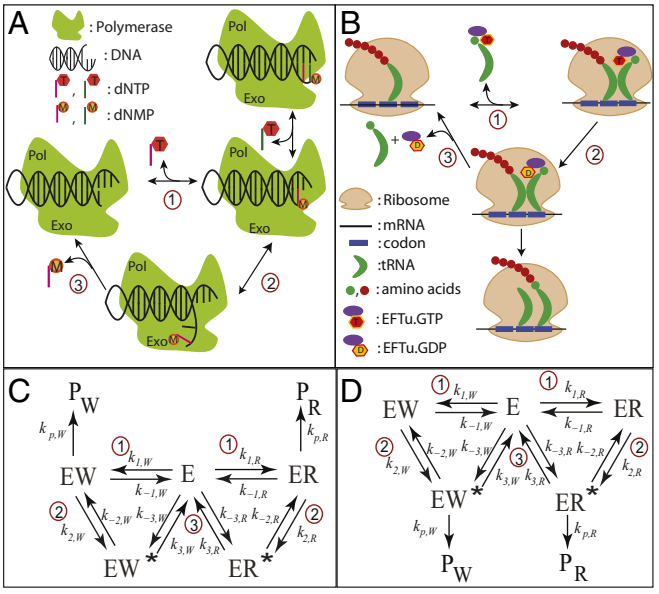

Fig. 1. Schematic representation of proofreading networks for (*A*) DNA replication by T7 DNAP enzyme and (*B*) aminoacyl(aa)-tRNA selection by *E. coli* ribosome during translation. Corresponding chemical networks are shown for (*C*) replication and (*D*) translation. Reaction steps comprising the cycles are labeled 1 to 3. Rate constants of each step are denoted by $k_{\pm i,\,\mathrm{R/W}}$, $i = 1, 2, 3$; subscript indicates right (R) or wrong (W) pathways. The rate constants of the steps leading to product (end) states are labeled as $k_{\mathrm{p,\,R/W}}$. The translation network in *D* is related to the replication network in *C* by the following transformation of rate constant indices: $\pm 1 \longleftrightarrow \mp 3$, $\pm 2 \longleftrightarrow \mp 2$. The steps involved in each case are, of course, different. For details, see *Methods*.

The schemes in Fig. 1 *A* and *C* are for DNA replication (3, 14). E denotes the T7 DNAP enzyme in complex with a DNA primer template. The R and W substrates are correct and incorrect base-paired dNTP molecules, respectively. Step 1 generates enzyme–DNA complexes ER(or EW) with the primer elongated by one nucleotide. Addition of another correct nucleotide to ER (EW) gives rise to P_R (P_W). ER* and EW* complexes denote the primer shifted to the exonuclease site (Exo) from the polymerase site (Pol) of DNAP. This commences proofreading in step 2. Excision of the nucleotide in step 3 resets the system to its initial state.

The schemes in Fig. 1 *B* and *D* show the aa-tRNA selection process by ribosome during translation (29). Here, E denotes the *E. coli* ribosome with mRNA. Cognate (near-cognate) aa-tRNAs in ternary complex with elongation factor Tu (EFTu) and GTP bind with ribosome in step 1 to form ER (EW). GTP hydrolysis in step 2 results in the complex ER* (EW*). The latter can take one of two routes. It can progress to the product P_R (P_W) with the elongation of the peptide chain by one amino acid. Alternatively, it can dissociate in the proofreading step (step 3), rejecting the aa-tRNA.

In both schemes, we take the rate constants of the W cycle to be related to those of the R cycle through $k_{\pm i,\,W} = f_{\pm i} k_{\pm i,\,R}$, $i = 1, 2, 3$ and similarly, for the catalytic step, $k_{p,\,W} = f_p k_{p,\,R}$. The set of rate constants $k_{1,\,R/W}$, $k_{-3,\,R/W}$ are effectively first order containing the substrate concentrations. The factors f_i provide the energetic discrimination between the R and W pathways. Completion of one cycle (returning to the starting state E) effectively amounts to hydrolysis of one dNTP molecule for DNA replication and one GTP molecule for aa-tRNA selection. The chemical potential difference, $\Delta\mu$ (in units of $k_B T$) is equal for both the cycles (19).

$$\Delta\mu = \ln\Big(\prod_{i=1}^{3} \frac{k_{i,\,R}}{k_{-i,\,R}}\Big) = \ln\Big(\prod_{i=1}^{3} \frac{k_{i,\,W}}{k_{-i,\,W}}\Big). \qquad [1]$$

This leads to the condition

$$\prod_{i=1,\,2,\,3} \frac{f_i}{f_{-i}} = 1. \qquad [2]$$

Accuracy and Speed from First-Passage Description. We determine the error and speed of the substrate selection kinetics from the first-passage probability density (27, 28). With this method, we can analyze an arbitrary catalytic reaction scheme and focus on the transitions starting from the initial state *E* that lead to the final state P_R. The description allows us to get analytical expression for both speed and accuracy for an arbitrary set of kinetic parameters. Therefore, we allow different rates for the R and W substrates for each step of the network as experimentally observed. Furthermore, we do not assume any step to be completely irreversible. Otherwise, the chemical potential difference over the cycle would diverge (Eq. **1**). This difference is linked to the hydrolysis of some energy-rich molecules supplying large but finite free energy.

Let us denote $F_{R,E}(t)$ as the probability density to reach state P_R at time t for the first time before reaching state P_W if the system is in state E at time $t = 0$. The corresponding probability density $F_{W,E}(t)$ is specified in the same manner. The evolution equations of $F_{R/W,E}(t)$ are known as the backward master equation (27, 30). It is more convenient to solve them in Laplace space (*Supporting Information*). We define the error, η as the ratio of the probabilities to reach the end states $P_{R/W}$ [also called the splitting probability (27)] given by

$$\eta = \frac{\Pi_W}{\Pi_R}; \quad \Pi_{R/W} = \int_0^\infty F_{R/W,E}(t)dt. \qquad [3]$$

It is important to note that this definition is equivalent to the traditional one (9, 21) defined as the ratio of the wrong product formation rate to the right one (see *Supporting Information*).

The speed of a reaction is naturally quantified by the net rate of the product formation. As with any chemical reaction rate, it can be defined as the inverse of the mean first-passage time (MFPT), i.e., the mean time it takes to cross the energy barrier that separates reactants and products for the first time. For example, a well-known application of this approach

for single-molecule MM kinetics results in the traditional expression for the rate as the inverse of the MFPT (31). We note that the speed toward the correct product can nevertheless be affected by the presence of the incorrect substrate. Thus, it is important to consider them together in contrast to the prevalent measure of the speed in literature neglecting the presence of the W pathway (21). In our case, the expression of the MFPT to reach each product state is given by the first moment of the corresponding probability density (27).

$$\tau_{R/W} = \frac{1}{\Pi_{R/W}} \int_0^\infty t\, F_{R/W,\,E}(t)dt. \quad [4]$$

In the *Results* and *Discussion*, we focus on τ_R as the measure of speed and denote it simply by τ.

Results

Although our formalism can be applied to an arbitrary KPR scheme, we've chosen to study two fundamentally important biologically processes: DNA replication and translation. These processes are best characterized in terms of underlying kinetic parameters, and we can study the speed–accuracy trade-off in the biologically relevant parameter region. Notably, despite differences in parameters and KPR mechanisms for the two case studies, we reach similar conclusions for both.

Importance of Speed over Accuracy in DNA Replication by T7 DNAP. The T7 DNAP enzyme catalyzes the polymerization of a DNA primer over a template strand (14). Wrongly incorporated dNTP is removed by the proofreading mechanism that involves the exonuclease site of DNAP (23). The model parameters of the corresponding reaction network (Fig. 1*A*) are listed in Table S1. They are based on the experimental data of Wong et al. (23). We do not consider dissociation of the DNA from the enzyme in our model. This approach is justified due to the faster polymerization rates in the R path and the higher exonucleolytic sliding rate in the W path (Table S1).

The error, η, varies among three limits as a function of the polymerization rate constant, $k_{1,R}(=k_{p,R})$ (23) with fixed f_1, f_p. All of them are lower bounds obtained in the limit $k_{1,R} \to 0$ (η_L), $k_{1,R} \to \infty$ (η_H), and an intermediate case with $k_{1,R} \lesssim k_{-1,R}$ (η_M). Explicit expressions follow from the general one for η (see *Supporting Information*). Here, we give suitable ratios of these limits to understand the error variation pattern

$$\frac{\eta_L}{\eta_M} = \frac{f_2}{f_{-1}}, \quad \frac{\eta_H}{\eta_M} = \frac{f_2}{f_p}\frac{k_{3,R}}{k_{-1,R}K_{M,R}}. \quad [5]$$

Here, $K_{M,R} = (k_{-2,R} + k_{3,R}/k_{2,R})$. From the experimental parameter values (Table S1) and Eq. **5**, we expect $\eta_M < \eta_L, \eta_H$. In other words, the system has a minimum error at some intermediate polymerization rate. This is indeed the case as shown in Fig. 2*A*. On the other hand, the MFPT, τ, decreases, and hence the speed increases monotonously with increase in $k_{1,R}$ (Fig. 2*B*). The range of τ also spans several orders of magnitude. The η–τ curve is shown in Fig. 2*C*. Negative slope of this curve indicates speed–accuracy trade-off, i.e., higher accuracy (lower η) corresponds to lower speed (higher τ) and vice versa. It is evident from Fig. 2*C* that there is a trade-off only when the polymerization rate constant becomes greater than the value corresponding to the minimum error. We call this branch with negative slope the trade-off branch. For lower values of $k_{1,R}$, error and MFPT change in the same direction. This branch with positive slope of the η–τ curve is denoted as the non-trade-off branch. Intuitively, the lack of trade-off for low polymerization rates arises due to different magnitudes of these rates between the R and W pathways; the latter has much smaller rates. When the polymerization rate is sufficiently smaller than the Pol–Exo sliding rate, correct substrate incorporation must undergo lots of unnecessary proofreading cycles. These futile cycles adversely affect the R pathway more, thereby compromising both speed and accuracy. Thus, at low polymerization rates, there is suboptimal regime allowing for improvement in both speed and accuracy.

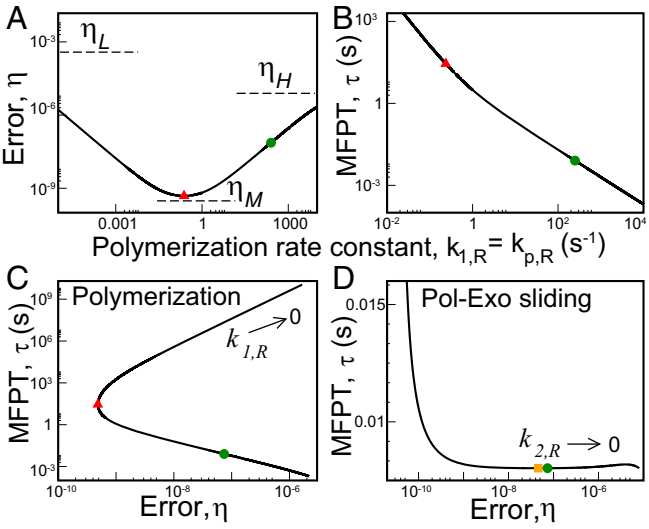

Fig. 2. Speed–accuracy trade-off for T7 DNAP. (*A*) The change in error, η, as a function of the polymerization rate constant $k_{1,R}(=k_{p,R})$. The error is bounded by the predicted limits (Eq. **5**). The green circle indicates the position of the actual system that is far away from the minimum error (red triangle). (*B*) Variation of MFPT, τ, with $k_{1,R}$. The red triangle gives the τ value corresponding to the minimum in η. (*C*) The η–τ curve for the polymerization step. (*D*) The η–τ curve for the Pol–Exo sliding step involved in proofreading generated by varying $k_{2,R}$ keeping f_2 fixed (semilog plot). There is a local minimum in τ (yellow square) near the actual value (green dot).

The actual system (green circle) is situated on the trade-off branch of the η–τ curve in Fig. 2*C*. It lies far away from the minimum error point (red triangle). In particular, the minimum error is ~150-fold lower than that of the actual system. However, to achieve this minimum error, the system's speed would drop by ~3,500-fold. Thus, the polymerization rate constant is selected to achieve high-enough speed. Significant amount of accuracy is lost in the process. The system can further lower the MFPT by moving down the slope of the η–τ curve. However, that means giving up more accuracy. Thus, of course, there is also a tolerable upper level of η.

The Pol–Exo sliding is an important step in error correction. The η–τ curve for this step is plotted in Fig. 2*D*. The minimum error value is approached in an asymptotic fashion at very large $k_{2,R}$. In contrast, the global minimum of MFPT is obtained in the $k_{2,R} \to 0$ limit. The MFPT also has a local minimum (yellow square) and a local maximum at finite $k_{2,R}$. Interestingly, the actual system lies pretty close to this (local) minimum. In particular, the system's τ value is almost identical to the minimum τ (within a less than 0.01%). On the other hand, the corresponding error η is ~1.6-fold higher than that corresponding to the minimum MFPT. The speed–accuracy trade-off appears after $k_{2,R}$ crosses the value corresponding to the local minimum in τ (and also before the local maximum). Thus the system is positioned on the non-trade-off branch of the η–τ curve. As one moves in either direction from the minimum τ point, error can change greatly with slight alteration in τ until η is too low. Therefore, speed appears to be more important as long as the system remains reasonably accurate. However, the system can gain lower errors at similar speeds by moving left to the trade-off branch of the η–τ curve. Then, what is the reason for not taking that route? We note that the proofreading pathway resets the system to the starting condition without progressing to product formation. Therefore, speedup in proofreading rate can increase the associated extra energy cost (11); this may restrict the system

to go to the more advantageous regime that has greater KPR rate and so, somewhat larger cost. We will further elaborate on this point in our next case study.

tRNA Selection by *E. coli* Ribosome Is Optimized for Speed Rather than Accuracy with a Cost Constraint. During translation, the ribosome decodes the mRNA sequences by selecting aa-tRNAs in ternary complex with EFTu and GTP (4, 15). Noncognate aa-tRNAs are removed by proofreading dissociation of the complex from the ribosome A site after GTP hydrolysis (24, 29). The model parameters of the network (Fig. 1*B*) for WT *E. coli* ribosome are listed in Table S2. They are based on the experimental data of Zaher and Green (29). We chose $k_{-2,R} = k_{-3,R} = 10^{-3}\ \mathrm{s}^{-1}$ to ensure that both step 2 and step 3 are nearly irreversible (29). There remains one free parameter f_{-2} (as f_{-3} gets fixed from Eq. **2**). We assumed $f_{-2} = 1$, but our main conclusions are independent of this choice (Fig. S1).

We show the η–τ curves for GTP hydrolysis and ternary complex binding steps in Fig. 3 *A* and *B*, respectively, for three varieties of *E. coli* ribosome. One is the WT, and the other two are mutants. One mutant, *rpsL141*, is hyperaccurate (HYP), and the other mutant, *rpsD12*, is more error-prone (ERR) than WT (29). Variation of the hydrolysis rate constant, $k_{2,R}$ (keeping f_2 fixed) results in quite large changes in error and MFPT. The trends are similar for all three systems (see Tables S3 and S4 for parameter sets of mutants). As for the polymerization steps in DNA replication, error varies among three bounds for the GTP hydrolysis step. They are obtained from the general expression of η (see *Supporting Information*) for low, intermediate, and high values of $k_{2,R}$,

$$\eta_{\mathrm{L}} = \alpha\frac{f_{-3}}{f_3}, \quad \eta_{\mathrm{M}} = \alpha\frac{f_1 f_2}{f_{-1} f_3}, \quad \eta_{\mathrm{H}} = \alpha\frac{f_1}{f_3}, \qquad \textbf{[6]}$$

where $\alpha = k_{p,W}/k_{3,R}$. For the parameter set of the system (Table S2, $f_{-2} = 1$), one gets $\eta_{\mathrm{M}} < \eta_{\mathrm{L}}, \eta_{\mathrm{H}}$. Thus, the error vs. hydrolysis rate curve passes through a minimum. Interestingly, there is also a minimum in τ as shown in Fig. 3*A*. The two minima are at different $k_{2,R}$ values, however. The speed–accuracy trade-off occurs between the minimum η (red triangle, WT) and the minimum τ (yellow square, WT) points. As is evident from Fig. 3*A*, all of the systems are positioned close to the minimum τ and far away from minimum η. For example, the WT ribosome would become ∼500-fold slower to achieve the minimum error, although the latter is ∼50-fold lower than the actual value. Hence, speed is preferred to accuracy. We tested the generality of this claim against multiple parameter variations (Fig. S2), and it appears

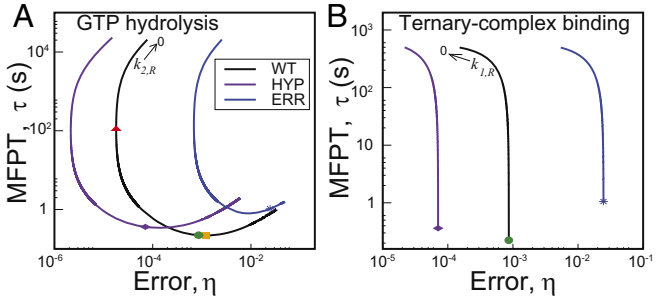

Fig. 3. Speed–accuracy trade-off in aa-tRNA selection by three varieties of *E. coli* ribosome. One is the wild-type (WT). The other two are hyperaccurate (HYP) and more error-prone (ERR) mutants. (*A*) η–τ curves for the GTP hydrolysis step. The actual system (green circle, WT) is situated close to the minimum τ (yellow square) and far away from minimum η (red triangle); this is also true for the mutants. (*B*) Speed–accuracy trade-off for the ternary complex binding step. The minimum error is achieved in the $k_{1,R} \to 0$ limit. MFPTs for all of the systems are close to saturation.

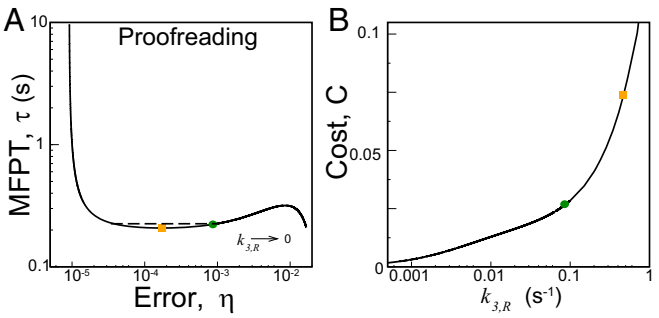

Fig. 4. (*A*) An η–τ diagram for the proofreading step. The actual system (green dot, WT) has a τ value similar to the local minimum in τ (yellow square). The dashed line shows the range up to which the system can lower the error with no loss in speed. (*B*) The proofreading cost, C, as a function of $k_{3,R}$. Its value at the local minimum in τ (yellow square) is approximately threefold higher than that of the actual system (green dot).

that speed is indeed more important. The robustness of this result is also tested successfully against fluctuations of the rate constants (Fig. S3). Interestingly, the WT ribosome is faster and, hence, better optimized for speed than both the mutants. It is important to note that the more accurate mutant HYP was not chosen by the natural selection. This point further emphasizes the importance of speed over accuracy in translation.

The change in the ternary complex binding rate constant, $k_{1,R}$ (with f_1 fixed) also affects both the error and the MFPT significantly, but on a smaller scale than hydrolysis (Fig. 3*B*). There is always a trade-off between speed and accuracy, unlike the cases studied so far. The minimum in error is obtained in the $k_{1,R} \to 0$ limit, whereas the maximum speed is achieved for very large $k_{1,R}$. With increase in $k_{1,R}$, τ falls several orders in magnitude. Interestingly, all of the systems have τ values almost identical to their respective saturation limits. To attain that state, they sacrifice an order of magnitude in terms of accuracy. Therefore, regarding speed–accuracy trade-off, the system is inclined to be faster with higher but tolerable error.

Next, we explore the effects of variation of the proofreading step rate constant, $k_{3,R}$ (keeping f_3 fixed) on system performance. The resulting η–τ diagram is plotted in Fig. 4*A* for the WT ribosome. The global minimum of τ is obtained in the $k_{3,R} \to 0$ limit, whereas the minimum of η lies in the large $k_{3,R}$ limit. There is a local minimum (yellow square) of τ at some intermediate $k_{3,R}$ along with a local maximum. Mutant ribosomes have similar trends (not shown in the figure). The nature of the η–τ curve is qualitatively similar to that obtained for the proofreading Pol–Exo sliding step in DNA replication (Fig. 2*D*) with two speed–accuracy trade-off branches. The actual system (green circle) is located on the non-trade-off branch that links the two trade-off branches of the η–τ curve. It has an MFPT close to (∼1.1-fold higher than) the (local) minimum τ value. However, the minimum τ point also has an approximately fivefold lower η. More important is the fact that the system can attain much lower errors with similar, even slightly higher, speeds if it moves left up to a certain level (the dashed line in Fig. 4*A*). What prevents the system from gaining in both speed and accuracy? Because correction by proofreading resets the system without a product formation, it has a cost associated with futile cycles where the correct substrate was inserted and then removed. The cost of proofreading, C, is defined as the ratio of the resetting flux to the product formation flux including both R and W pathways (11) (see *Supporting Information*); this gives a measure of the amount of extra energy-rich molecules consumed due to the presence of the proofreading step. Specifically, the cost C can quantify the moles of dNTP (or GTP) used for proofreading per mole of product (11). This quantity can be easily computed from

our formalism and investigated as a function of kinetic parameters. In particular, we quantify how the cost of proofreading changes with the increase in $k_{3,R}$ near local minimum of the η–τ curve in Fig. 4*A*. The results, shown in Fig. 4*B*, demonstrate that the cost associated with the (local) minimum τ point is approximately threefold higher than that of the actual system. This cost disadvantage (threefold higher GTP consumption per amino acid!) may restrict the system from gaining the available advantage in both speed and accuracy. A similar consideration may also be responsible for the nature of trade-off exhibited in the DNA replication case (Fig. 2*D*).

Discussion

Evolution has optimized the kinetic parameters of biological enzymes to achieve the desired levels of accuracy and speed at various stages of biological information flow. In this study, by examining how the balance between speed and accuracy changes with variation of the underlying kinetic parameters, we gain insights into the important priorities for this optimization. To this end, we focus on two fundamental examples of biological proofreading networks: DNA replication and protein translation. In both cases, the systems tend to achieve maximum speed by losing significant accuracy. However, the speed–accuracy trade-off only occurs in the limited region of the parameter space, e.g., after the polymerization rate in replication passes the minimum error point. In the case of translation, the trade-off appears between the minima in error and the MFPT for the GTP hydrolysis step. A similar conclusion about the importance of speed over accuracy is reached by varying the rates of the proofreading steps in both systems. Although higher proofreading rates can further improve the accuracy without losing much speed, the associated energy cost of proofreading may restrict further improvements on an already acceptable speed and accuracy.

An important insight from the above analyses is that the speed–accuracy trade-off is not universally present, and its occurrence depends on the specific values of kinetic rates. Biologically, this result implies that mutations or application of drugs that reduce the enzyme's accuracy do not necessarily increase its speed, and vice versa. The widespread view of a compromise between accuracy and speed is mainly based on their dependence on the effective catalytic rate of the process (9, 21). Indeed, the larger the catalytic rate, the higher the speed and the lower the accuracy. However, the role of other steps, like hydrolysis and proofreading, are not as straightforward. Our study reveals that, for these steps, trade-offs are present only over a certain range of rates, and both accuracy and speed can improve with variation of certain kinetic parameters. The partitioning of the error–time curves into trade-off and non-trade-off branches clarifies the distinct roles of various transitions and the molecular mechanisms of the speed–accuracy optimization. Our conclusions are also supported by a more advanced analysis of the maximum speed vs. accuracy curves using Pareto fronts, as explained in detail in *Supporting Information*.

The analysis of speed–accuracy trade-offs for different mutant varieties of *E. coli* ribosome further confirms the importance of speed over accuracy. The WT and two mutants (HYP and ERR) lie close to the minimum MFPT point on the error–time curves (Fig. 3*A*). However, the WT and HYP ribosomes are on the trade-off branch, whereas the ERR mutant is on the non-trade-off branch. Thus, movement down the slope toward the trade-off branch would raise both accuracy and speed for the ERR ribosome. That is how the WT ribosome may have evolved from the more erroneous ERR type. However, any further movement upward along the trade-off branch means a slowdown with a lower error; this leads to the more accurate (HYP) mutant. Rejection of the latter as the natural choice implies that optimization of speed is critical. We note that comparison of *E. coli* growth rates with WT and mutant ribosomes already indicates such an optimization (21, 32). However, according to the prevailing notion on the ever-present compromise between error and speed, the more erroneous (ERR) ribosome should be faster. Hence, the hindered growth for ERR mutant was ascribed presumably to less-active proteins (33). Our results indicate that not only the accuracy but also the speed of peptide chain elongation can be smaller for the ERR mutant.

Despite different schemes and parameter values of the replication and translation networks, there appears to be a general mechanism of error correction; this becomes apparent from the trade-off diagrams for the proofreading step. A rate constant of the proofreading step in both the cases is selected such that speed of the system is close to the maximum possible one. The actual systems reside on the non-trade-off branch of their respective error–time curves. Biologically, that implies that mutation that slightly speeds up the proofreading step would lead to an increase in both speed and accuracy of the enzymes. However, we show that such mutation would also increase energetic costs of proofreading. This extra cost does not allow the systems to further reduce the error and MFPT. Furthermore, the most interesting feature for both of the systems is the proximity of the MFPT value to the local minimum, which is similar in magnitude to the global minimum. Hence, for both case studies, the KPR rate is fine-tuned so that the loss in speed is insignificant compared with the improvement in accuracy.

Our results on the accuracy–speed trade-off in two important biological networks reveal similar strategies to optimize these two vital quantities. Rates of the steps like substrate binding, hydrolysis (of intermediates), and catalysis seem to be chosen to enhance speed at the cost of accuracy. On the other hand, proofreading or error-correction steps seem to be selected to have such rates that the error is reduced sufficiently with almost no loss in speed. Therefore, between the maximization of accuracy and speed, biological systems appear to give precedence to the latter. Tolerable levels of error and cost of error correction act as constraints to tailor the speed. It is interesting to note here that experimentally observed distribution of discriminatory steps is not optimal from the point of view of minimizing error (34). For example, for ribosome, the rates of the catalytic step are significantly different between the incorporation of the R and W amino acid in the polypeptide chain. Although this may be suboptimal in terms of error minimization (34), it allows for the proofreading rate to be much faster than the catalytic rate for a W substrate and much slower than the catalytic rate for the R substrate. As a result, ribosome avoids futile cycles (correcting the errors it did not make), improving speed and energy cost. This observation gives additional support to our arguments that biological systems distribute discrimination to better optimize speed and not accuracy (see *Supporting Information*). Our study thus presents a coherent quantitative picture of how the ultimate balance between accuracy and speed is achieved by adjusting various rates in distinct ways. It will be important to test our predictions in other systems and organisms. We believe such testing will further help to elucidate the fundamental mechanisms of proofreading processes in biological systems.

ACKNOWLEDGMENTS. This work is supported by Center for Theoretical Biological Physics National Science Foundation (NSF) Grant PHY-1427654. A.B.K. also acknowledges support from Welch Foundation Grant C-1559 and from NSF Grant CHE-1360979.

1. Alberts B, et al. (2013) *Essential Cell Biology* (Garland Sci, New York), 4th Ed.
2. Watson JD (1976) *Molecular Biology of the Gene* (WA Benjamin, New York).
3. Kunkel TA, Bebenek K (2000) DNA replication fidelity. *Annu Rev Biochem* 69: 497–529.

4. Zaher HS, Green R (2009) Fidelity at the molecular level: Lessons from protein synthesis. *Cell* 136:746–762.
5. Reynolds NM, Lazazzera BA, Ibba M (2010) Cellular mechanisms that control mistranslation. *Nat Rev Microbiol* 8:849–856.
6. Nangle LA, de Crecy Lagard V, Doring V, Schimmel P (2002) Genetic code ambiguity cell viability related to severity of editing defects in mutant tRNA synthetases. *J Biol Chem* 277:45729–45733.
7. Schimmel P (2008) Development of tRNA synthetases and connection to genetic code and disease. *Protein Sci* 17:1643–1652.
8. Pauling L (1957) The probability of errors in the process of synthesis of protein molecules. *Festschrift Arthur Stoll* (Birkhauser, Basel), pp 597–602.
9. Hopfield JJ (1974) Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci USA* 71:4135–4139.
10. Ninio J (1975) Kinetic amplification of enzyme discrimination. *Biochimie* 57:587–595.
11. Freter RR, Savageau MA (1980) Proofreading systems of multiple stages for improved accuracy of biological discrimination. *J Theor Biol* 85:99–123.
12. Ehrenberg M, Blomberg C (1980) Thermodynamic constraints on kinetic proofreading in biosynthetic pathways. *Biophys J* 31:333–358.
13. Hopfield JJ, Yamane T, Yue V, Coutts SM (1976) Direct experimental evidence for kinetic proofreading in amino acylation of tRNAIle. *Proc Natl Acad Sci USA* 73: 1164–1168.
14. Johnson KA (1993) Conformational coupling in DNA polymerase fidelity. *Annu Rev Biochem* 62:685–713.
15. Gromadski KB, Rodnina MV (2004) Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell* 13:191–200.
16. Blanchard SC, Gonzalez RL, Kim HD, Chu S, Puglisi JD (2004) tRNA selection and kinetic proofreading in translation. *Nat Struct Mol Biol* 11:1008–1014.
17. Kunkel TA (2004) DNA replication fidelity. *J Biol Chem* 279:16895–16898.
18. Murugan A, Huse DA, Leibler S (2012) Speed, dissipation, and error in kinetic proofreading. *Proc Natl Acad Sci USA* 109:12034–12039.
19. Hartich D, Barato AC, Seifert U (2015) Nonequilibrium sensing and its analogy to kinetic proofreading. *New J Phys* 17:055026.
20. Thompson RC, Karim AM (1982) The accuracy of protein biosynthesis is limited by its speed: High fidelity selection by ribosomes of aminoacyl-tRNA ternary complexes containing GTP [gamma S]. *Proc Natl Acad Sci USA* 79:4922–4926.
21. Johansson M, Lovmar M, Ehrenberg M (2008) Rate and accuracy of bacterial protein synthesis revisited. *Curr Opin Microbiol* 11:141–147.
22. Johansson M, Zhang J, Ehrenberg M (2012) Genetic code translation displays a linear trade-off between efficiency and accuracy of tRNA selection. *Proc Natl Acad Sci USA* 109:131–136.
23. Wong I, Patel SS, Johnson KA (1991) An induced-fit kinetic mechanism for DNA replication fidelity: Direct measurement by single-turnover kinetics. *Biochemistry* 30: 526–537.
24. Wohlgemuth I, Pohl C, Mittelstaet J, Konevega AL, Rodnina MV (2011) Evolutionary optimization of speed and accuracy of decoding on the ribosome. *Phil Trans R Soc B* 366:2979–2986.
25. Sartori P, Pigolotti S (2013) Kinetic versus energetic discrimination in biological copying. *Phys Rev Lett* 110:188101.
26. Savir Y, Tlusty T (2013) The ribosome as an optimal decoder: A lesson in molecular recognition. *Cell* 153:471–479.
27. van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam), 3rd Ed.
28. Bel G, Munsky B, Nemenman I (2010) The simplicity of completion time distributions for common complex biochemical processes. *Phys Biol* 7:016003.
29. Zaher HS, Green R (2010) Hyperaccurate and error-prone ribosomes exploit distinct mechanisms during tRNA selection. *Mol Cell* 39:110–120.
30. Kolomeisky AB (2015) *Motor Proteins and Molecular Motors* (CRC, Boca Raton, FL).
31. Kou SC, Cherayil BJ, Min W, English BP, Xie XS (2005) Single-molecule Michaelis - Menten equations. *J Phys Chem B* 109:19068–19081.
32. Tubulekas I, Hughes D (1993) Suppression of rpsL phenotypes by tuf mutations reveals a unique relationship between translation elongation and growth rate. *Mol Microbiol* 7:275–284.
33. Lovmar M, Ehrenberg M (2006) Rate, accuracy and cost of ribosomes in bacterial cells. *Biochimie* 88:951–961.
34. Murugan A, Huse DA, Leibler S (2014) Discriminatory proofreading regimes in nonequilibrium systems. *Phys Rev X* 4:021016.