Trade-Offs between Error, Speed, Noise, and Energy Dissipation in Biological Processes with Proofreading

Joel D. Mallory,[†] Anatoly B. Kolomeisky,[*,†,‡,§,‖] and Oleg A. Igoshin[*,†,‡,⊥]

[†]Center for Theoretical Biological Physics, [‡]Department of Chemistry, [§]Department of Chemical and Biomolecular Engineering, [‖]Department of Physics and Astronomy, and [⊥]Department of Bioengineering and Department of Biosciences, Rice University, Houston, Texas 77005, United States

Ⓢ *Supporting Information*

ABSTRACT: High accuracy of major biological processes relies on the ability of the participating enzymatic molecules to preferentially select the correct substrate from a pool of chemically similar substrates by activating the so-called proofreading mechanisms. While the importance of such mechanisms is widely accepted, it is still unclear how evolution has optimized the biological systems with respect to certain characteristic properties. Here, using a discrete-state stochastic framework with a first-passage analysis, we theoretically investigate trade-offs between four characteristic properties of enzymatic systems, namely, error, speed, noise, and energy dissipation. Specifically, two fundamental biological processes are examined, i.e., DNA replication in the T7 bacteriophage and tRNA selection during protein translation in *Escherichia coli*. Notably, all of the characteristic properties cannot be completely optimized at the same time due to trade-offs between them. To understand the relative importance of the computed quantities to the enzymatic functionality, we introduce a new quantitative metric to rank the properties. The results demonstrate that the reaction speed is the principal characteristic property that evolution optimizes in both enzymatic systems and that the energy dissipation comes in second. In addition, the error and the noise are always ranked third and fourth, respectively, regardless of the system considered. Physicochemical arguments to explain these observations are presented.
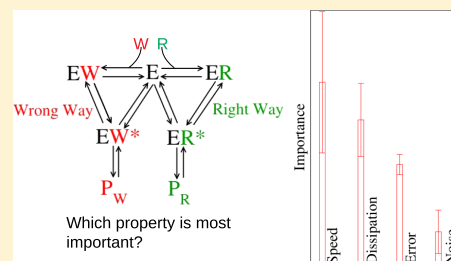
INTRODUCTION

Fundamental biological processes in living cells such as DNA replication, RNA transcription, and protein translation are known to be highly accurate.[1] This stems from the catalytic activity of enzymatic molecules that exhibit remarkable reaction speeds and striking substrate selectivity.[2,3] The error rates for RNA transcription and protein translation are on the order of $\eta \sim 10^{-4}-10^{-5}$ and $10^{-3}-10^{-4}$,[4,5] respectively, while the error rate for DNA replication is even lower at $\eta \sim 10^{-8}-10^{-10}$.[2] It is believed that the high fidelity of these major biological processes is due to the presence of various proofreading mechanisms.[2,6−10] The most famous example of such error-correction processes is the kinetic proofreading (KPR) mechanism that was independently proposed and explained by Hopfield and Ninio.[11,12] KPR enables an enzyme to discriminate between correct (cognate) and wrong (noncognate) substrates while also detecting and removing wrongly incorporated substrates by transitioning the system back to its original state with an additional proofreading step. This mechanism and its implications for biological systems have been intensively studied,[4,13−21] and the existence of the KPR mechanism in various biological systems is supported by multiple experimental observations.[4,14,19,22−25]

Analysis of the KPR mechanism suggests that it enhances the accuracy of the enzymatic process by resetting the system back to the initial state to correct the error without proceeding to the product state. However, this action simultaneously reduces the overall speed of the process, which may constitute a physiologically unfavorable price to pay for living cells. Hence, it was argued that biological systems must find a proper compromise between speed, i.e., the inverse of the mean first-passage time (MFPT) and the accuracy to optimize their functionality. To this end, a recent theoretical investigation performed an analysis of trade-offs between speed and accuracy in enzymatic networks for the *Escherichia coli* ribosome and T7 DNA polymerase enzymes.[9] These characteristic properties were explicitly examined in the presence of both right and wrong kinetic pathways.[9] It was discovered that a speed−accuracy trade-off does indeed exist in some situations and that the speed was determined to be more important than the accuracy for both enzymes. Thus, previous theoretical research on these enzymes revealed that evolution preferred to keep the overall enzymatic speed as high as possible while still maintaining the error at tolerable levels.[9] However, at the same time, the error and the MFPT were shown to exhibit nontrivial, i.e., nonmonotonic, behavior upon the variation of some kinetic parameters, and the speed−accuracy trade-off was not always present.[9]

Although the speed and the error are important properties to analyze the performance of enzymatic processes, they are by no

means the only characteristic properties that biological systems can optimize. For example, the proofreading step in the KPR mechanism is frequently associated with futile cycles and the expenditure of energy-rich nucleotide triphosphate (NTP) molecules,[9] which means that the proofreading step requires free-energy consumption.[26−29] Thus, it seems plausible to suggest that biological systems will aim to keep the energy dissipation small enough such that it does not become detrimental to enzymatic functionality. Recent theoretical results on the proofreading cost for the *E. coli* ribosome and T7 DNA polymerase lend support to this argument as a notable increase in the energy expenditure from proofreading prevented evolution from completely optimizing both the speed and accuracy of the enzymes.[9] Another important characteristic property of biological processes is the noise, which is the deviation of the dynamic properties of the system (e.g., the MFPT) from the average values due to the stochastic nature of the underlying biochemical reactions.[30] The consistency in the overall reaction speeds of such biological processes suggests that they also attempt to limit the noise levels. At the present time, it remains unclear whether biological systems are optimized with respect to these characteristic properties and to what extent. Despite recent theoretical studies on relations between these properties,[21] their relative importance is generally unknown.

In this work, trade-offs between the error, speed, noise, and energy dissipation in enzymatic networks with KPR mechanisms are theoretically investigated. Using a recently developed theoretical method based on a first-passage analysis of discrete-state stochastic networks,[9,10] the importance of the four characteristic properties is examined for the *E. coli* ribosome and T7 DNA polymerase. For these enzymes, experimental data are available and provide a comprehensive chemical kinetic description.[14,19,23,24] The ribosome catalyzes the process of aminoacyl(aa)-tRNA selection (i.e., polypeptide chain elongation) during protein translation,[3] while the polymerase carries out the process of DNA replication.[2] We have determined that evolution has fine-tuned the rate constants for both enzymes such that the characteristic properties are within their optimal regimes, but in certain cases, trade-offs between the properties prevent all of them from becoming completely optimized. Note that here and throughout this work, we refer to the minimization or maximization of the characteristic properties as criteria for optimization. In addition, a new quantitative metric to determine the relative importance of the criteria is developed and tested on both systems. Our theoretical results provide a more coherent molecular picture of how complex biological systems function so accurately and efficiently.

■ METHODS

Chemical Kinetic Proofreading Networks for the *E. coli* Ribosome and T7 DNA Polymerase. We provide an in-depth discussion of the chemical kinetic networks for the *E. coli* ribosome and T7 DNA polymerase enzymes. Although these representations are approximate models of the true biochemical network of states, they are composed of several important steps on the enzymatic pathways for which the rate constants have been determined experimentally (see Figure 1).[14,19,23,24] Notably, our descriptions of each chemical kinetic network include proofreading and product formation cycles for both the right R and wrong W substrates, and they are based
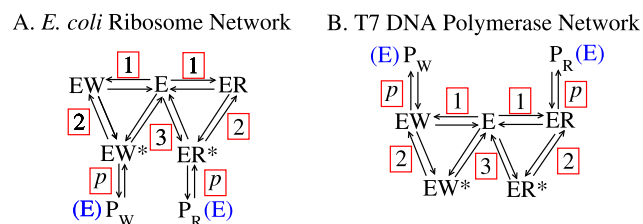
A. *E. coli* Ribosome Network   B. T7 DNA Polymerase Network

Figure 1. Chemical kinetic proofreading networks for aminoacyl-tRNA selection in protein translation by the *E. coli* ribosome and for DNA replication by the T7 DNA polymerase considered in this work. The labels for each step on the kinetic pathways correspond to the different rate constants $k_{\pm i, R/W}$, $i = 1, 2, 3, p$ for the right R and for the wrong W pathways of the kinetic schemes. The KPR networks are composed of two complete cycles for each type of substrate R or W: proofreading and product formation. The proofreading cycles involve steps 1, 2, and 3, while the product formation cycles are composed of steps 1, 2, and p for the ribosome, and steps 1 and p for the polymerase. Note that the states denoted as $P_{R/W}$ are actually variations of the free enzyme state (i.e., the blue state (E)) such that the transitions labeled p constitute the final steps of the product formation cycles.

on the KPR mechanism introduced independently by Hopfield and Ninio.[11,12]

The chemical kinetic network for the *E. coli* ribosome is presented in Figure 1A. The free ribosome in state E binds to an mRNA transcript and translates the genetic information encoded within triplets of nucleotides, i.e., codons, into a correct sequence of amino acids. Charged cognate (right R) or noncognate (wrong W) aa-tRNA molecules bearing complimentary anticodons, elongation factor Tu (EFTu), and GTP enter the ribosome in step 1 to form enzymatic complexes, which are denoted as ER (right pathway) or EW (wrong pathway) in Figure 1A. In step 2, GTP hydrolysis is taking place, and the process continues to states ER* or EW*, respectively. These are also enzyme−substrate complexes, but they are in different conformational states, now with GDP bound instead of GTP. At this stage, the ribosome can add one more amino acid via peptide bond formation to the growing polypeptide chain by taking the pathway denoted p to achieve the final product states P_R or P_W (see Figure 1A), which constitute a complete product formation cycle. Then, the ribosome can continue adding more amino acids to the nascent polypeptide chain starting again from the state (E). However, from the states ER* or EW*, there is a possibility for the system to return to the initial free enzyme state E without completing the product formation cycle. If the ribosome encounters a noncognate aa-tRNA (i.e., the codon and the anticodon do not correctly match and the system is in the state EW*), the probability of activating the proofreading pathway is higher.[4,19]

The chemical kinetic scheme for the T7 DNA polymerase enzyme from bacteriophage T7 is slightly different from the ribosome, and it is shown in Figure 1B. This enzyme is responsible for the replication of a primer DNA strand. The polymerase functions as a molecular motor that hydrolyzes cognate (right R) or noncognate (wrong W) deoxyNTP molecules and then links them together via phosphodiester bonds in step 1 such that the polymerase proceeds from the free enzyme state E to ER (right pathway) or EW (wrong pathway) (see Figure 1B). From here, the polymerase can take one of two routes to complete the DNA replication process. If the enzyme takes the pathway labeled p, a second nucleotide is

added to the nascent DNA strand and the final products P_R or P_W are generated, which completes the product formation cycle. The polymerase then continues the processes starting again from state (E). Given that the correct substrate was incorporated, the new nucleotides are complementary to those on the original DNA primer–template. In the second route, an exonuclease (Exo) domain of the enzyme is activated in step 2 such that the enzyme advances to states ER* or EW* (see Figure 1B). Finally, the Exo domain removes the nucleotide in step 3, and the system is restored to the initial free enzyme state E. Experiments suggest that the probability of the polymerase taking the proofreading route is higher when it accepts a wrong nucleotide in step 1.[14,23]

In our theoretical analysis, we utilize first-order (or pseudo-first-order) chemical kinetic rate constants $k_{\pm i,R}$ determined experimentally for the *E. coli* ribosome[19] and T7 DNA polymerase,[14,23,24] and they are presented in Table 1. The

Table 1. Experimental Rate Constants $k_{\pm i,R}$ and Discrimination Factors $f_{\pm i}$ for the *E. coli* Ribosome and the T7 DNA Polymerase[a]

System Parameters	*E. coli* Ribosome	T7 DNA Polymerase
$k_{1,R}$	40	250
$k_{-1,R}$	0.5	1
$k_{2,R}$	25	0.2
$k_{-2,R}$	1×10^{-3}	700
$k_{3,R}$	8.5×10^{-2}	900
$k_{p,R}$	8.415	250
f_1	0.675	8×10^{-6}
f_{-1}	94	1×10^{-5}
f_2	4.8×10^{-2}	11.5
f_{-2}	1	1
f_3	7.9	1
f_p	4.2×10^{-3}	4.8×10^{-5}

[a]The values of $k_{\pm i,R}$ (in units of s^{-1}) and $f_{\pm i}$ (dimensionless) for the ribosome are from ref 19, and those for the polymerase are from refs 14, 23, 24. Note that the value of $k_{-2,R}$ was reported in ref 9. All of the rate constants are first-order (or pseudo-first-order) in units of s^{-1}. The rate constants $k_{-3,R}$ and $k_{-p,R}$ were determined from the broken detailed balance condition using fixed chemical potential differences $\Delta\mu_{NTP}$ and $\Delta\mu_p$ (see the text) that were set to their physiological values (see eq 1). The discrimination factors f_{-3} and f_{-p} were also determined from a secondary constraint (see eq 2).

remaining rate constants ($k_{-3,R}$ and $k_{-p,R}$) not shown in Table 1 are determined from the broken detailed balance conditions by fixing the chemical potential differences at their physiological values such that $\Delta\mu_{NTP} \sim 20\ k_BT$ for the KPR cycles, $\Delta\mu_p \sim 26\ k_BT$ for peptide bond formation in the ribosome,[31–33] and $\sim 11\ k_BT$ for phosphodiester bond formation. Note from Figure 1A,B that every biochemical transition in the KPR reaction networks is reversible to avoid unphysical situations with diverging chemical potential differences. Moreover, we imposed a second constraint on the broken detailed balance conditions for $\Delta\mu_{NTP}$ and $\Delta\mu_p$ such that the chemical potential differences for the R and W cycles are equivalent according to

$$\Delta\mu_{NTP,p} = \ln\left(\prod_{i=1}^{N}\frac{k_{i,R}}{k_{-i,R}}\right) = \ln\left(\prod_{i=1}^{N}\frac{k_{i,W}}{k_{-i,W}}\right) \quad (1)$$

Note that the total number of biochemical states on the reaction networks is $N = 3$ for the proofreading cycles and $N =$

p for the product formation cycles (see Figure 1A,B). Table 1 also includes discrimination factors $f_{\pm i}$ for all of the steps on the reaction pathways. The discrimination factors are defined in the following way: the rate constants for the R and W pathways, $k_{\pm i,R}$ and $k_{\pm i,W}$, are related to each other via $f_{\pm i} = k_{\pm i,W}/k_{\pm i,R}$. The values of the discrimination factors f_{-3} and f_{-p} not shown in Table 1 are determined from the following relation

$$\prod_{i=1}^{N}\frac{f_i}{f_{-i}} = 1 \quad (2)$$

which is required by the first constraint in eq 1. Note that upon variation of a single rate constant $k_{i,R}$, the corresponding discrimination factor f_i remains fixed.

Equations for the Error, MFPT, Noise, and Normalized Energy Dissipation. To quantify the degree of optimization of the four criteria for the *E. coli* ribosome and T7 DNA polymerase, analytic equations for the error η, enzymatic turnover time (MFPT) τ, and noise δ_τ are derived using a first-passage analysis of the discrete-state stochastic systems.[34] One can introduce first-passage probability density functions $F_{R/W,E}(t)$ to reach the right product state P_R at time t before reaching the wrong one P_W for the first time given that the system started in the free enzyme state E at time $t = 0$. It can be shown that all of the dynamic properties of the system can be expressed in terms of these functions. The temporal evolution of first-passage probabilities is governed by a set of backward master equations.[34] Solving this system via a Laplace transformation gives the explicit expressions for $F_{R/W,E}(t)$. Then, all characteristic properties of the system can be easily evaluated. For example, the so-called splitting probabilities $\Pi_{R/W}$ for reaching the end states P_R and P_W at all times are given by

$$\Pi_{R/W} = \int_0^\infty F_{R/W,E}(t)\mathrm{d}t \quad (3)$$

Then, the error rate η is defined as the dimensionless ratio of the splitting probability for reaching the wrong end state Π_W to the one for reaching the right end state Π_R[9]

$$\eta = \frac{\Pi_W}{\Pi_R} \quad (4)$$

Another important quantity for our analysis is the conditional MFPT $\tau \equiv \langle t \rangle$, which is equal to the inverse enzymatic rate for reaching the right product state (see Figure 1A,B). It is the mean time (in seconds) required for the enzyme to reach the state P_R before reaching the state P_W for the first time. In the first-passage framework, the MFPT is defined as the first moment of the first-passage probability density $F_{R,E}(t)$ such that

$$\tau \equiv \langle t \rangle = \frac{1}{\Pi_R}\int_0^\infty tF_{R,E}(t)\mathrm{d}t \quad (5)$$

The second moment of $F_{R,E}(t)$ is defined in a similar fashion as

$$\langle t^2 \rangle = \frac{1}{\Pi_R}\int_0^\infty t^2F_{R,E}(t)\mathrm{d}t \quad (6)$$

Using eqs 5 and 6, an expression for the noise δ_τ is obtained as the dimensionless ratio of the standard deviation $\sqrt{\langle t^2 \rangle - \langle t \rangle^2}$ and the MFPT

$$\delta_\tau = \frac{\sqrt{\langle t^2 \rangle - \langle t \rangle^2}}{\langle t \rangle} \tag{7}$$

In other words, the noise in eq 7 can be interpreted as the coefficient of variation for the first-passage time.[30]

The total energy dissipation σ (i.e., the entropy production in units of $k_B T$) is the rate at which an enzyme (or, in principle, any nonequilibrium system) generates heat.[26] The KPR reaction networks for the ribosome and the polymerase (see Figure 1A,B) have proofreading and product formation cycles (involving both the right R and the wrong W substrates) with corresponding chemical potential differences $\Delta\mu_{NTP}$ for NTP hydrolysis and $\Delta\mu_p$ for product formation from eq 1, i.e., the cycle affinities. Therefore, the total energy dissipation over all of the cycles can be written as

$$\sigma = J_{proof}\Delta\mu_{NTP} + J_p\Delta\mu_p \tag{8}$$

where J_{proof} and J_p are the respective steady-state proofreading and production formation fluxes, which can be determined using a theoretical framework prescribed by Koza.[35] Dividing eq 8 by the rate of NTP hydrolysis per molecule of the product formed, i.e., $\sigma_N = \sigma/J_p\Delta\mu_{NTP}$ gives a dimensionless quantity σ_N

$$\sigma_N = \frac{J_{proof}}{J_p} + \frac{\Delta\mu_p}{\Delta\mu_{NTP}} \tag{9}$$

Thus, eq 9 defines σ_N as the total, normalized energy dissipation for the two KPR networks studied in this work, which depends on both the kinetic cost of proofreading and the ratio of the thermodynamic cost of product formation to that of NTP hydrolysis. Moreover, eq 9 constitutes a more robust and more convenient definition of the proofreading cost in comparison to a phenomenological quantity introduced earlier, i.e., J_{proof}/J_p.[9] In the Supporting Information, more detailed derivations of the equations shown here for all four of the characteristic properties are presented. The relative importance of the characteristic properties for the *E. coli* ribosome and T7 DNA polymerase is directly assessed using these equations.

■ RESULTS AND DISCUSSION

We begin by analyzing trade-offs between the characteristic properties in the *E. coli* ribosome and T7 DNA polymerase. The trade-off plots (see, e.g., Figure 2) are generated by varying a single rate constant $k_{i,R}$ to see how evolution changes the characteristic properties while keeping all of the other kinetic parameters fixed at their experimentally determined values (see Table 1). If an increase in one of the characteristic properties corresponds to a decrease in the other, there is a trade-off between them, and the native system is located on a branch of the trade-off plot with a negative slope. This indicates that evolution cannot optimize both properties simultaneously. In contrast, if the characteristic properties both increase and decrease together, the native system is on a non-trade-off branch with a positive slope. This implies that evolution might be able to optimize both properties at the same time provided that this action does not detrimentally affect the other characteristic properties.

Furthermore, we seek not only to analyze the characteristic properties in terms of the trade-offs between them but also to ascertain the relative importance of the properties as criteria for optimization. For each rate constant k_i corresponding to a
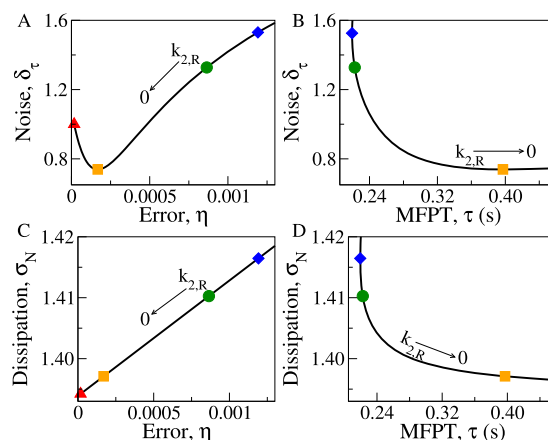

Figure 2. Trade-off plots demonstrating the interplay between four pairs of characteristic properties studied in this work due to variation of the rate constant for GTP hydrolysis $k_{2,R}$ in the *E. coli* ribosome. (A) $\delta_\tau - \eta$; (B) $\delta_\tau - \tau$; (C) $\sigma_N - \eta$; and (D) $\sigma_N - \tau$. The green dot denotes the native system, while the red triangles, orange squares, and blue diamonds represent the positions of the minimum error, noise, and MFPT, respectively.

selected step on the chemical kinetic network, we determine how close its native value k_0 is to the position of a local minimum for the given characteristic property $f(k_i)$ ($f = \eta, \tau, \delta_\tau, \sigma_N$). In this way, we denote the value of k_i at which the minimum (maximum) occurs as k_{min} (k_{max}). To eliminate any constant terms that could bias our results, we normalize this expression by the difference between $f(k_{min})$ and $f(k_{max})$ and then take the $\log_{10}$ to define the metric of the scaled difference

$$d_{k,i} \equiv -\log_{10}\left(\frac{f(k_0) - f(k_{min})}{f(k_{max}) - f(k_{min})}\right) \tag{10}$$

which quantifies the degree of optimization for each criterion. Note that a value of $d_{k,i} = 1$ for the scaled difference corresponds to an order of magnitude variation by 10^{-1} on the natural scale and so forth. In addition, the different rate constants k_i are varied within a predetermined range $k_i \in [\alpha, \beta]$. The interval is centered around the native rate constant k_0 (see Table 1) such that the interval bounds are $\alpha = 10^{-3}k_0$ and $\beta = 10^3 k_0$, respectively. Using eq 10, we ranked the importance of the four criteria shown in Figure 2 according to their $d_{k,i}$ values. Criteria that have larger $d_{k,i}$ values are ranked higher than criteria with smaller $d_{k,i}$ values.

***E. coli* Ribosome: Trade-Offs between the Characteristic Properties.** Trade-off plots created by varying the rate constant for GTP hydrolysis $k_{2,R}$ in the ribosome are presented in Figure 2. One can see that the native system (green dot) is on non-trade-off branches of the noise−error and dissipation−error plots where there is a positive slope (Figure 2A,C). At first glance, it would appear that the native system could evolve to decrease the noise by ~80% (and by extension, also the error and the dissipation) by sliding further down the non-trade-off branch to the minimum noise (orange square). However, this scenario is prevented by a ~78% increase in the MFPT, i.e., ~44% loss of speed. In addition, there are also noise−MFPT and dissipation−MFPT trade-offs such that the native system is on branches of the trade-off plots with negative slopes (see Figure 2B,D). Note that the MFPT of the native system could still decrease by ~1.4% such that the ribosome would achieve the minimum MFPT (blue diamond).

However, it seems that the evolution did not enforce complete optimization of the MFPT as this would result in ~38% increase in the error, ~15% increase in the noise, and ~0.4% increase in the dissipation. Applying the scaled difference $d_{k,i}$ shows that when the rate constant for GTP hydrolysis $k_{2,R}$ is varied, the MFPT has the largest $d_{k,2R} = 4.11$ from eq 10. However, the rest of the characteristic properties have $d_{k,2R}$ values separated from that for the MFPT by a difference of at least ~2.5. Thus, our results suggest that the corresponding rate constant $k_{2,R}$ is probably "tuned" by the evolutionary process in such a way that the overall reaction speed is the most important criterion for this particular step on the enzymatic pathway.

A similar analysis can be performed for rate constants corresponding to other steps on the chemical kinetic pathway for the ribosome. Considering the rate constant for proofreading $k_{3,R}$, we note that the noise for the native system (green dot) is located on non-trade-off branches of the noise–error and noise–MFPT plots (see Figure 3A,B). If evolution
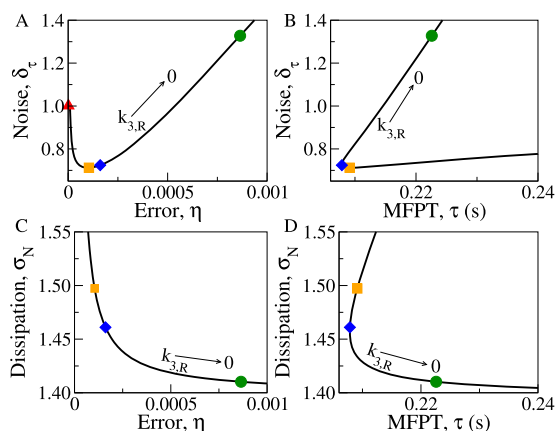

Figure 3. Trade-off plots demonstrating the interplay between four pairs of characteristic properties studied in this work due to variation of the rate constant for proofreading $k_{3,R}$ in the *E. coli* ribosome. (A) $\delta_\tau - \eta$; (B) $\delta_\tau - \tau$; (C) $\sigma_N - \eta$; and (D) $\sigma_N - \tau$. The green dot denotes the native system, while the red triangles, orange squares, and blue diamonds represent the positions of the minimum error, noise, and MFPT, respectively.

attempted to optimize the noise by ~86% to its minimum value (orange square), there would be a resulting ~6.1% increase in the dissipation. Furthermore, there are dissipation–error and dissipation–MFPT trade-offs as well (Figure 3C,D). It seems that decreasing the MFPT by ~7.1% to its minimum value (blue diamond) is avoided as this would cause a ~3.6% increase in the dissipation. The dissipation has the largest scaled difference $d_{k,3R} = 2.58$, and the $d_{k,3R}$ values for the other characteristic properties are separated from that for the dissipation by a gap of least ~0.9. As such, the rate constant for proofreading $k_{3,R}$ is optimized to keep the dissipation tolerable. This explanation seems reasonable because the ribosome must not waste too much energy provided by GTP hydrolysis when proofreading is activated on the chemical kinetic pathway.

Finally, the variation of the rate constant for product formation $k_{p,R}$ reveals that the native system (green dot) is on a non-trade-off branch of the noise–error trade-off plot (see Figure 4A). It would seem as though evolution could push the native system closer to the local minimum value (orange
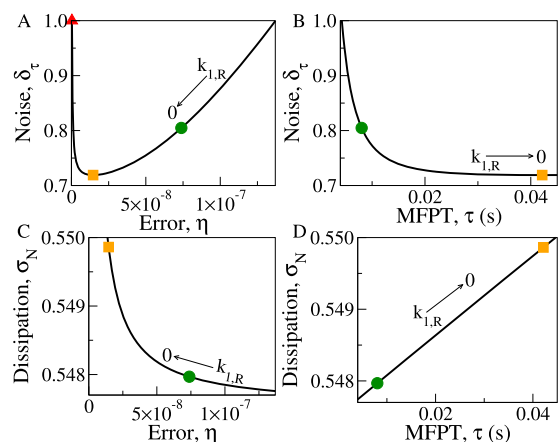

Figure 4. Trade-off plots demonstrating the interplay between four pairs of characteristic properties studied in this work due to variation of the rate constant for product formation $k_{p,R}$ in the *E. coli* ribosome. (A) $\delta_\tau - \eta$; (B) $\delta_\tau - \tau$; (C) $\sigma_N - \eta$; and (D) $\sigma_N - \tau$. The green dot denotes the native system, while the red triangles and orange squares and blue diamonds represent the positions of the minimum error and noise, and MFPT respectively.

square) in the noise, thereby decreasing this value by ~40%. However, this would not be advantageous for the ribosome as this action would result in a ~4-fold increase in the MFPT and a ~4% increase in the dissipation (cf. Figure 4B–D). The rankings of the criteria from eq 10 place the MFPT and the dissipation on roughly equal footing with $d_{k,pR} = 2.92$ and 2.59, respectively. The other properties have $d_{k,pR}$ values that are smaller than those corresponding to the MFPT and the dissipation by a gap of at least ~1.3. Apparently, the evolutionary process optimized the rate constant for product formation $k_{p,R}$ such that the final step of polypeptide chain elongation can occur quickly without wasting too much energy in the process.

T7 DNA Polymerase: Trade-Offs between the Characteristic Properties. In a similar manner, we can analyze the trade-offs between the characteristic properties for the T7 DNA polymerase. Let us start by varying the rate constant for polymerization $k_{1,R} = k_{p,R}$ with results presented in Figure 5. The native system (green dot) is on a non-trade-off branch of the noise–error trade-off plot (see Figure 5A). Apparently, evolution did not allow the polymerase to decrease the noise by ~12% and simultaneously to lower the error by ~5-fold (orange square) as this would result in a ~4-fold increase in the MFPT (i.e., ~81% loss in speed) due to the trade-off between the noise and the MFPT (Figure 5B). In addition, decreasing the noise to achieve the minimum value would create other problems for the polymerase as the dissipation would increase slightly by ~0.3% (Figure 5C,D). Rankings of the criteria from eq 10 indicate that the MFPT has the largest scaled difference $d_{k,1R} = 3.51$ followed by the dissipation with $d_{k,1R} = 3.01$. The remaining characteristic properties have $d_{k,1R}$ values that are smaller than the MFPT and the dissipation by a difference of ~0.8. Therefore, the evolutionary process optimized the rate constant for polymerization $k_{1,R} = k_{p,R}$ to make the speed and dissipation the most important criteria for this step on the DNA replication network.

Moving to the variation of the rate constant for Pol–Exo sliding $k_{2,R}$ for proofreading, the native system (green dot) is located on non-trade-off branches of the noise–error and noise–MFPT parametric plots (Figure 6A,B). If the noise
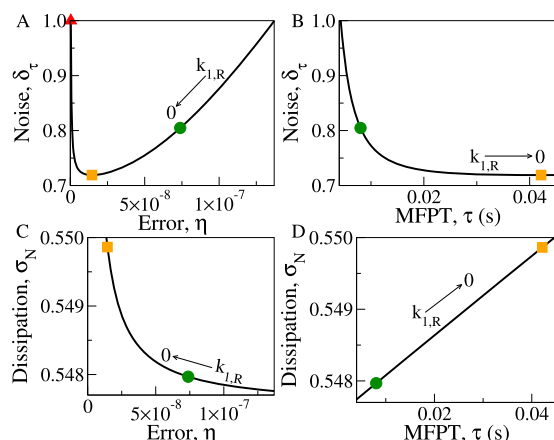
Figure 5. Trade-off plots demonstrating the interplay between four pairs of characteristic properties studied in this work due to variation of the rate constant for polymerization $k_{1,R} = k_{p,R}$ in the T7 DNA polymerase. (A) $\delta_\tau - \eta$; (B) $\delta_\tau - \tau$; (C) $\sigma_N - \eta$; and (D) $\sigma_N - \tau$. The green dot denotes the native system, while the red triangles and orange squares represent the positions of the minimum error and noise, respectively.
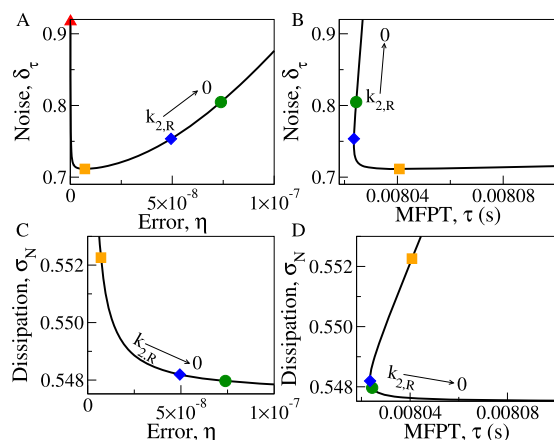

Figure 6. Trade-off plots demonstrating the interplay between four pairs of characteristic properties studied in this work due to variation of the rate constant for Pol−Exo sliding $k_{2,R}$ in the T7 DNA polymerase. (A) $\delta_\tau - \eta$; (B) $\delta_\tau - \tau$; (C) $\sigma_N - \eta$; and (D) $\sigma_N - \tau$. The green dot denotes the native system, while the red triangles, orange squares, and blue diamonds represent the positions of the minimum error, noise, and MFPT, respectively.

could be decreased by ∼13% to its minimum value (orange square), the situation would only be less beneficial to the functionality of the polymerase as this would cause a ∼1% increase in the dissipation and a ∼0.2% increase in the MFPT. The native system is also on trade-off branches of the dissipation−error and dissipation−MFPT trade-off plots (Figure 6C,D). Further decreasing the MFPT by ∼0.01% to its optimal value (blue diamond) is prevented as this would result in a small ∼0.05% increase in the dissipation due to the trade-off between these two characteristic properties. Note also that the trade-off plots in Figure 6 are shown in the Supporting Information (see Figure S1) with a larger range of variation for the MFPT and energy dissipation. The rankings of the criteria using the parameter $d_{k,2R}$ demonstrate that the MFPT has the largest scaled difference $d_{k,2R} = 3.42$ followed by the dissipation at $d_{k,2R} = 2.99$. Moreover, the noise and the error have $d_{k,2R}$ values that are similar to the dissipation such that $d_{k,2R} = 2.20$

and 1.99 for these criteria, respectively. Thus, the evolutionary process optimized the rate constant for Pol−Exo sliding $k_{2,R}$ such that this step is fast, does not expend too much energy, and proceeds without many unnecessary fluctuations or errors. This scenario makes sense for the Pol−Exo sliding step as the polymerase should not have to waste unnecessary time or energy switching between the polymerase and exonuclease domains attempting to fix the errors it made.

Overall Rankings of the Criteria for Optimization. Our theoretical approach enables us to determine overall rankings and hence the degrees of optimization for the four criteria (i.e., error η, MFPT τ, noise δ_τ, and energy dissipation σ_N). Composite averages D_{mean} for the ribosome and the polymerase are shown as bar plots in Figure 7A,B, respectively. The
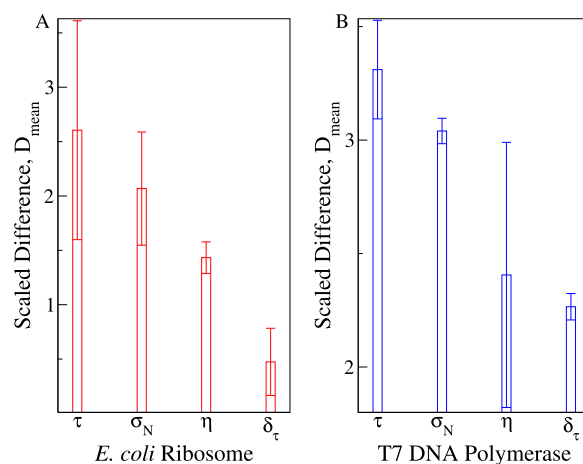

Figure 7. Composite averages D_{mean} of the scaled differences $d_{k,i}$ displayed as bar plots (note the scales) for each of the four criteria considered in this work (i.e., error η, MFPT τ, noise δ_τ, and energy dissipation σ_N). Note that the range of D_{mean} values from the MFPT to the noise is wider for the ribosome ((A) red bars) at ∼2.1, while it is more narrow for the polymerase ((B) blue bars) at ∼1.0.

sets of scaled differences are such that $d_{k,i} \in \mathcal{D}$ are composed of various $d_{k,i}$ values (see Supporting Information Figure S2) for all of the forward rate constants $k_{i,R}$ shown previously and one additional reverse rate constant $k_{-1,R}$. As such, the total number of $d_{k,i}$ values N in each set of scaled differences is $N = 4$ for the ribosome and $N = 3$ for the polymerase. The average values corresponding to the different criteria

$$D_{mean} = \frac{1}{N} \sum_{i=1}^{N} d_{k,i} \tag{11}$$

are shown in Figure 7 for each enzyme. The overall rankings of the criteria are determined using eq 11 in the same way as for the individual $d_{k,i}$ values. The criterion with the largest D_{mean} value receives the highest rank, while the criterion with the smallest D_{mean} value is assigned the lowest rank. The corresponding standard deviations (i.e., the error bars reported in Figure 7) were obtained using the following equation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (d_{k,i} - D_{mean})^2}{N}} \tag{12}$$

For the ribosome, the MFPT has the largest $D_{mean} = 2.60$ followed by the dissipation $D_{mean} = 2.07$, indicating that the MFPT has the highest rank with the dissipation coming in

second (see Figure 7A). Likewise, the MFPT and the dissipation follow the same order for the polymerase such that the MFPT is the most important criterion with D_{mean} = 3.31, while the dissipation finishes in second place with D_{mean} = 3.04 (see Figure 7B). In contrast, the error and the noise are always ranked third and fourth for both enzymes, respectively, making these two criteria the least important. The averages with the largest and the smallest D_{mean} values (i.e., the MFPT and the noise) are separated over a wider range of ~2.1 for the ribosome, but for the polymerase, this gap is more narrow (i.e., a factor of ~2 smaller) at ~1.0. In addition, the gap between the MFPT and the dissipation values is larger for the ribosome at ~0.54 than it is for the polymerase at ~0.27. Thus, the smaller difference in D_{mean} values for the MFPT and the dissipation in the polymerase indicates that these two criteria are closer in terms of relative importance than in the ribosome. The larger range of D_{mean} values for the ribosome demonstrates that evolution had a stronger preference to minimize the MFPT (i.e., maximize the overall reaction speed), while the other three criteria do not matter as much. However, the criteria are more comparable for the polymerase as the D_{mean} values are found within a smaller range, but the overall reaction speed still prevails as the most important criterion.

CONCLUSIONS

We theoretically investigated the trade-offs between four characteristic properties in essential biological processes, i.e., the error, speed (i.e., the inverse of the MFPT), noise, and energy dissipation. Our approach employs a discrete-state stochastic method with a first-passage framework that provides analytic equations for all of the properties considered. We applied this formalism to the *E. coli* ribosome and the T7 DNA polymerase enzymes, for which experimental data are available. We find that the characteristic properties are nearly optimized, but the existence of trade-offs between them prevents evolution from completely optimizing all of them at the same time. A new quantitative metric to rank the importance of the four criteria for optimization is introduced. Our results demonstrate that the overall reaction speed is the most important criterion in both enzymatic networks and it is followed by the energy dissipation. The error and the noise are less important criteria, and they are always ranked third and fourth, respectively. Although these qualitative trends are the same for both biological systems examined, there are some quantitative differences that reflect how evolution preferentially optimized the criteria for each enzyme.

It is also important to critically evaluate our theoretical approach. We note that the characteristic properties analyzed here are not all independent from each other (i.e., the equations describing the properties η, τ, and δ_τ have some terms in common). As such, the rankings determined from the scaled difference $d_{k,i}$ within a specified range of rate constants may be affected by this dependence. However, expanding the range of rate constants by 2 orders of magnitude around the native value k_0 (i.e., using an interval of $k_i \in [10^{-4}k_0, 10^4 k_0]$ as shown in the Supporting Information Figure S3) does not alter the overall rankings of the criteria, which indicates that our ranking method is robust. Moreover, we derived a more complete expression for the normalized energy dissipation (see eq 9) than the phenomenological one used previously. This equation depends not only on the proofreading cost[9] but also on the ratio of the thermodynamic cost of product formation to that of NTP hydrolysis. Furthermore, we took into account

only a few of the most important biochemical states comprising the chemical kinetic networks and therefore the networks should be regarded as approximations of the true, nontrivial biological systems. This raises the question regarding the extent to which our predictions will be affected by accounting for more intermediate biochemical states. In addition, we only addressed four possible criteria, but other characteristic properties may also be optimized to support the functionality of biological systems. It would also be interesting to apply the mathematical formalism and the ranking method to other enzymatic systems that display high fidelity (i.e., low error rates) due to the presence of the KPR mechanism, e.g., aminoacyl-tRNA synthetase.[22,36,37]

ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.9b03757.

> Derivation of the analytic equations for the error η, MFPT τ, and noise δ_τ; derivation of the analytic equation for the normalized energy dissipation σ_N; additional trade-off plots for the variation of the rate constant for Pol−Exo sliding $k_{2,R}$ with extended ranges of the axes as compared to those in Figure 6; bar plots of the individual $d_{k,i}$ values for the four criteria used to determine the composite averages D_{mean} in Figure 7; composite averages D_{mean} determined using a larger interval of rate constants than in Figure 7 (PDF)

(PDF)

AUTHOR INFORMATION

Corresponding Authors
*E-mail: tolya@rice.edu (A.B.K.).
*E-mail: igoshin@rice.edu (O.A.I.).
ORCID ⓘ
Anatoly B. Kolomeisky: 0000-0001-5677-6690
Oleg A. Igoshin: 0000-0002-1449-4772
Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Alberts, B.; Bray, D.; Hopkin, K.; Johnson, A. D.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Essential Cell Biology*; Garland Science, Taylor and Francis Group: New York, NY, 2014.
(2) Kunkel, T. A. DNA Replication Fidelity. *J. Biol. Chem.* **2004**, *279*, 16895−16898.
(3) Ninio, J. Multiple Stages in Codon-Anticodon Recognition: Double-Trigger Mechanisms and Geometric Constraints. *Biochimie* **2006**, *88*, 963−992.
(4) Zaher, H. S.; Green, R. Fidelity at the Molecular Level: Lessons from Protein Synthesis. *Cell* **2009**, *136*, 746−762.
(5) Reynolds, N. M.; Lazazzera, B. A.; Ibba, M. Cellular Mechanisms that Control Mistranslation. *Nat. Rev. Microbiol.* **2010**, *8*, 849−856.

(6) Johnson, K. A. Conformational Coupling in DNA Polymerase Fidelity. *Annu. Rev. Biochem.* **1993**, *62*, 685−713.

(7) Wohlgemuth, I.; Pohl, C.; Mittelstaet, J.; Konevega, A. L.; Rodnina, M. V. Evolutionary Optimization of Speed and Accuracy of Decoding on the Ribosome. *Philos. Trans. R. Soc., B* **2011**, *366*, 2979−2986.

(8) Gromadski, K. B.; Rodnina, M. V. Kinetic Determinants of High-Fidelity tRNA Discrimination on the Ribosome. *Mol. Cell* **2004**, *13*, 191−200.

(9) Banerjee, K.; Kolomeisky, A. B.; Igoshin, O. A. Elucidating Interplay of Speed and Accuracy in Biological Error Correction. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 5183−5188.

(10) Banerjee, K.; Kolomeisky, A. B.; Igoshin, O. A. Accuracy of Substrate Selection by Enzymes is Controlled by Kinetic Discrimination. *J. Phys. Chem. Lett.* **2017**, *8*, 1552−1556.

(11) Hopfield, J. J. Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes requiring High Specificity. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4135−4139.

(12) Ninio, J. Kinetic Amplification of Enzyme Discrimination. *Biochimie* **1975**, *57*, 587−595.

(13) Bennett, C. H. Dissipation-Error Tradeoff in Proofreading. *BioSystems* **1979**, *11*, 85−91.

(14) Wong, I.; Patel, S. S.; Johnson, K. A. An Induced-Fit Kinetic Mechanism for DNA Replication Fidelity: Direct Measurement by Single-Turnover Kinetics. *Biochemistry* **1991**, *30*, 526−537.

(15) Murugan, A.; Huse, D. A.; Leibler, S. Speed, Dissipation, and Error in Kinetic Proofreading. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 12034−12039.

(16) Murugan, A.; Huse, D. A.; Leibler, S. Discriminatory Proofreading Regimes in Nonequilibrium Systems. *Phys. Rev. X* **2014**, *4*, No. 021016.

(17) Barato, A. C.; Seifert, U. Thermodynamic Uncertainty Relation for Biomolecular Processes. *Phys. Rev. Lett.* **2015**, *114*, No. 158101.

(18) Barato, A. C.; Seifert, U. Universal Bound on the Fano Factor in Enzyme Linetics. *J. Phys. Chem. B* **2015**, *119*, 6555−6561.

(19) Zaher, H. S.; Green, R. Hyperaccurate and Error-Prone Ribosomes exploit Distinct Mechanisms during tRNA Selection. *Mol. Cell* **2010**, *39*, 110−120.

(20) Hartich, D.; Barato, A. C.; Seifert, U. Nonequilibrium Sensing and its Analogy to Kinetic Proofreading. *New J. Phys.* **2015**, *17*, No. 055026.

(21) Wong, F.; Amir, A.; Gunawardena, J. Energy-Speed-Accuracy Relation in Complex Networks for Biological Discrimination. *Phys. Rev. E* **2018**, *98*, No. 012420.

(22) Hopfield, J.; Yamane, T.; Yue, V.; Coutts, S. Direct Experimental Evidence for Kinetic Proofreading in Amino Acylation of tRNA[Ile]. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 1164−1168.

(23) Donlin, M. J.; Patel, S. S.; Johnson, K. A. Kinetic Partitioning between the Exonuclease and Polymerase Sites in DNA Error Correction. *Biochemistry* **1991**, *30*, 538−546.

(24) Patel, S. S.; Wong, I.; Johnson, K. A. Pre-Steady-State Kinetic Analysis of Processive DNA Replication including Complete Characterization of an Exonuclease-Deficient Mutant. *Biochemistry* **1991**, *30*, 511−525.

(25) Jakubowski, H. Proofreading in *trans* by an Aminoacyl-tRNA Synthetase: A Model for Single Site Editing by Isoleucyl-tRNA Synthetase. *Nucleic Acids Res.* **1996**, *24*, 2505−2510.

(26) Qian, H.; Kjelstrup, S.; Kolomeisky, A. B.; Bedeaux, D. Entropy Production in Mesoscopic Stochastic Thermodynamics: Nonequilibrium Kinetic Cycles Driven by Chemical Potentials, Temperatures, and Mechanical Forces. *J. Phys: Condens. Matter* **2016**, *28*, No. 153004.

(27) Pigolotti, S.; Neri, I.; Roldán, É.; Jülicher, F. Generic Properties of Stochastic Entropy Production. *Phys. Rev. Lett.* **2017**, *119*, No. 140604.

(28) Gingrich, T. R.; Horowitz, J. M.; Perunov, N.; England, J. L. Dissipation Bounds All Steady-State Current Fluctuations. *Phys. Rev. Lett.* **2016**, *116*, No. 120601.

(29) Hyeon, C.; Hwang, W. Physical Insight into the Thermodynamic Uncertainty Relation using Brownian Motion in Tilted Periodic Potentials. *Phys. Rev. E* **2017**, *96*, No. 012156.

(30) Bel, G.; Munsky, B.; Nemenman, I. The Simplicity of Completion Time Distributions for Common Complex Biochemical Processes. *Phys. Biol.* **2009**, *7*, No. 016003.

(31) Jakubowski, H.; Goldman, E. Quantities of Individual Aminoacyl-tRNA Families and their Turnover in *Escherichia coli*. *J. Bacteriol.* **1984**, *158*, 769−776.

(32) Bremer, H.; Dennis, P. P. *Escherichia coli and Salmonella: Cellular and Molecular Biology*; ASM Press: Washington, D.C., 1996.

(33) Acosta-Silva, C.; Bertran, J.; Branchadell, V.; Oliva, A. Quantum-Mechanical Study on the Mechanism of Peptide Bond Formation in the Ribosome. *J. Am. Chem. Soc.* **2012**, *134*, 5817−5831.

(34) Kolomeisky, A. B. *Motor Proteins and Molecular Motors*; CRC Press, Taylor and Francis Group: Boca Raton, FL, 2015.

(35) Koza, Z. General Technique of Calculating the Drift Velocity and Diffusion Coefficient in Arbitrary Periodic Systems. *J. Phys. A: Math. Gen.* **1999**, *32*, 7637−7651.

(36) Santra, M.; Bagchi, B. Catalysis of tRNA Aminoacylation: Single Turnover to Steady-State Kinetics of tRNA Synthetases. *J. Phys. Chem. B* **2012**, *116*, 11809−11817.

(37) Santra, M.; Bagchi, B. Kinetic Proofreading at Single Molecular Level: Aminoacylation of tRNA[Ile] and the role of water as an editor. *PLoS One* **2013**, *8*, No. e66112.