The Journal of Chemical Physics

The guiding role of dissipation in kinetic proofreading networks: Implications for protein synthesis

Cite as: J. Chem. Phys. **152**, 111102 (2020); https://doi.org/10.1063/1.5144726
Submitted: 10 January 2020 • Accepted: 01 March 2020 • Published Online: 17 March 2020

Kinshuk Banerjee, Biswajit Das and Gautam Gangopadhyay

The Journal of Chemical Physics COMMUNICATION scitation.org/journal/jcp

The guiding role of dissipation in kinetic proofreading networks: Implications for protein synthesis

Cite as: J. Chem. Phys. 152, 111102 (2020); doi: 10.1063/1.5144726
Submitted: 10 January 2020 • Accepted: 1 March 2020 •
Published Online: 17 March 2020

Kinshuk Banerjee,[1] Biswajit Das,[2] and Gautam Gangopadhyay[2,a)]

AFFILIATIONS

[1] Department of Chemistry, Acharya Jagadish Chandra Bose College, Kolkata 700 020, India
[2] S. N. Bose National Centre for Basic Sciences, Block JD, Sector III, Salt Lake City, Kolkata 700 106, India

[a)] **Author to whom correspondence should be addressed:** gautam@bose.res.in

ABSTRACT

Major biological polymerization processes achieve remarkable accuracy while operating out of thermodynamic equilibrium by utilizing the mechanism known as kinetic proofreading. Here, we study the interplay of the thermodynamic and kinetic aspects of proofreading by exploring the dissipation and catalytic rate, respectively, under the realistic constraint of fixed chemical potential difference. Theoretical analyses reveal *no-monotonic* variations of the catalytic rate and total entropy production rate (EPR), the latter quantifying the dissipation, at steady state. Applying this finding to a tRNA selection network in protein synthesis, we observe that the network tends to *maximize* both the EPR and catalytic rate, but not the accuracy. Simultaneously, the system tries to minimize the ratio of the EPRs due to the proofreading steps and the catalytic steps. Therefore, dissipation plays a guiding role in the optimization of the catalytic rate in the tRNA selection network of protein synthesis.

I. INTRODUCTION

An open system, such as any living organism, exchanges mass and energy with the environment and works out of thermodynamic equilibrium.[1] A characteristic feature of such a system is a positive value of the total entropy production rate (EPR)[2,3] even when the system reaches steady state, designated generally as the non-equilibrium steady state (NESS).[4,5] In NESS, the total EPR multiplied by the temperature gives the dissipation.[6] To sustain such a NESS, the presence of a driving force is essential. In chemical and biochemical reaction networks, this force is the finite chemical potential difference.[4,7,8] In the presence of a driving force, the reaction fluxes are non-zero even in NESS, whereas in equilibrium, such fluxes vanish as the force, i.e., the chemical potential difference, becomes zero.[1] Thus, the total EPR, which is a product of the fluxes and the corresponding forces,[2] also disappears in equilibrium.

Interestingly, an advantage of being out of equilibrium is the achievement of high accuracy in selecting the right substrate over the wrong ones in all the major biological polymerization processes such as DNA replication[9] and protein translation.[10] If the underlying reaction networks are assumed to operate in equilibrium, such high accuracy cannot be explained. The necessity of the non-equilibrium condition in explaining the experimentally determined accuracy data was proposed in the seminal works of Hopfield[11] and Ninio[12] in terms of the mechanism called kinetic proofreading (KPR).[13] Actually, Ninio described his error-correction mechanism as kinetic amplification that goes beyond the classical view of enzyme selectivity based on a "lock-and-key" picture. KPR is based on (i) driving the system out of equilibrium by coupling specific reaction steps to energy sources such as adenosine triphosphate (ATP) [guanosine triphosphate (GTP)] hydrolysis and (ii) resetting the system back to the initial state in a cyclic manner instead of going forward to form products. This allows the system to rectify the mistakes incurred due to the insertion of wrong substrates and, hence, attain low values of error, i.e., high accuracy, at the cost of energy dissipation.[14,15] However, the works of Hopfield and Ninio were based on irreversible biochemical reactions that precluded the

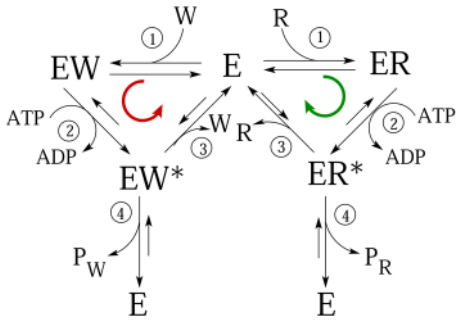

FIG. 1. Schematic diagram of the reaction network. The enzyme (E) can form complexes with the right (R) as well as the wrong (W) substrate leading to the right (P_R) and wrong product (P_W), respectively.

quantitative determination of entropy production (or free energy dissipation). This issue was addressed in a later work by Qian[16] using cycle networks with reversible steps. Now, the KPR mechanism explains the high accuracy with energy expenditure but also predicts a trade-off[11,17] between the accuracy and catalytic rate, i.e., higher accuracy leads to a lower rate of product formation. On the other hand, there are works on the co-polymerization model of DNA replication that show the opposite result.[18,19] These two contrasting results are reconciled in a later work by Sartori and Pigolotti.[20] A recent study, however, showed that the KPR networks involved in DNA replication and tRNA selection during protein synthesis are optimized to achieve a maximum rate rather than maximum accuracy.[21]

With the background discussed above, here, we investigate the non-equilibrium thermodynamic and kinetic aspects of proofreading, focusing on the total EPR and the catalytic rate. Our study is based on the network shown in Fig. 1, which is a generalization of the original Hopfield model[11] with the rate constants being different for each step between the right and the wrong pathways. This feature is motivated by experimental data on tRNA selection during protein synthesis that involve KPR.[10] Through the cycles comprising steps-1, -2, and -3 (see Fig. 1), the system resets via proofreading (step-3) after enzyme–substrate binding (step-1) and subsequent ATP hydrolysis (step-2). We call these cycles the resetting cycles (green for right pathway and red for wrong pathway). The net chemical potential difference over the resetting cycles is given by ATP hydrolysis only.[4,21] Therefore, *the chemical potential difference over the resetting cycles is kept fixed*.[22] The catalytic cycles, on the other hand, consist of steps 1, 2, and 4 where the catalysis steps are designated as step-4. Unlike the resetting cycles, the chemical potential differences over the catalytic cycles are different for right and wrong pathways. This is because, along with ATP and adenosine diphosphate (ADP), different substrates and products are involved in these cycles. Hence, we do not fix the chemical potential differences over the catalytic cycles. We explore the variations of the total EPR and the catalytic rate under the realistic biological constraint of fixed chemical potential difference in the light of the following questions: (i) Whether and how these two quantities are optimized? (ii) What are the roles of individual steps such as proofreading and catalysis in the optimization, if any?

II. NON-MONOTONIC VARIATIONS OF EPR AND CATALYTIC RATE UNDER FIXED CHEMICAL POTENTIAL DIFFERENCE

We start our theoretical investigation by expressing the chemical potential difference, $\Delta\mu_{123}$, over both the resetting cycles in terms of the rate constants of the KPR network as[21]

$$\frac{\Delta\mu_{123}}{k_BT} = \ln\left(\frac{k_{1,R}k_{2,R}k_{3,R}}{k_{-1,R}k_{-2,R}k_{-3,R}}\right) = \ln\left(\frac{k_{1,W}k_{2,W}k_{3,W}}{k_{-1,W}k_{-2,W}k_{-3,W}}\right). \quad (1)$$

Here, rate constants for the i-th step ($i = 1, \ldots, 4$) are denoted by $k_{\pm iR(W)}$, where "±"-sign indicates forward and corresponding reverse reactions, respectively. $k_{1R(W)}$ and $k_{-3R(W)}$ are pseudo-first-order rate constants containing R(W) substrate concentrations.[21]

The total EPR (in units of k_B), denoted simply as EPR from now on, for the respective cycles and for the whole network at NESS is given by[1,2]

$$\sigma_{R(W)} = \frac{1}{T}\sum_{i=1}^{4} J_{iR(W)}\Delta\mu_{iR(W)}; \ \sigma = \sigma_R + \sigma_W. \quad (2)$$

Here, $J_{iR(W)}$ is the flux and $\Delta\mu_{iR(W)}$ is the chemical potential difference for step-i of the R(W) pathway; T denotes the temperature of the system. The fluxes are interrelated at NESS by

$$J_{1R(W)} = J_{2R(W)} = J_{3R(W)} + J_{4R(W)}. \quad (3)$$

The catalytic rates are defined as $v_{R(W)} = k_{4R(W)}\tilde{P}_{ER^*(EW^*)}$, where $\tilde{P}_j$ denotes the NESS probability to remain in the state-j(=E, ER, ER^*, EW, EW^*). The expressions of the NESS fluxes, the corresponding forces, and state probabilities are provided in the supplementary material. The discrimination factors between the right and wrong paths are introduced as $\delta_i = \frac{k_{iW}}{k_{iR}}$, $(i = \pm1, \ldots, \pm4)$. The ratio of the catalytic rates of right and wrong pathways gives the accuracy, ϵ, of the KPR process,[13]

$$\epsilon = \frac{v_R}{v_W} = \frac{k_{4R}\tilde{P}_{ER^*}}{k_{4W}\tilde{P}_{EW^*}} = \frac{1}{\delta_4}\frac{\tilde{P}_{ER^*}}{\tilde{P}_{EW^*}}. \quad (4)$$

In order to keep $\Delta\mu_{123}$ constant, one needs to change more than one rate constant simultaneously, as suggested by Eq. (1). We maintain this constant by varying the pair of rate constants by keeping either the product or the ratio fixed. Previous studies suggest that the dissociation of the intermediates $ER(EW)$ and $ER^*(EW^*)$ and ATP hydrolysis play major roles in governing the accuracy and rate of KPR.[11,21] We study the variation of the EPR by keeping (i) $k_{2R}k_{3R}$ fixed and (ii) k_{3R}/k_{-1R} fixed for a fixed set of δ_i ($i = \pm1, \ldots, \pm4$) with $\Pi_{i=1}^{3}\frac{\delta_i}{\delta_{-i}} = 1$ following from Eq. (1).

Experimental data for real KPR networks suggest that except step-1, the rest of the steps are strongly irreversible.[23] Hence, we neglect the terms with $k_{-2R(W)}$, $k_{-3R(W)}$, and $k_{-4R(W)}$ in our theoretical analyses wherever possible. With this approximation, the NESS flux for step-1 reduces to

$$J_{1R} = k_{1R}\tilde{P}_E\left(1 - \frac{k_{-1,R}}{k_{-1,R} + k_{2,R}}\right), \quad (5)$$

with the details given in the supplementary material. First, we vary k_{2R} keeping $k_{2R}k_{3R}$ (also δ_is) fixed. At low k_{2R}, the quantity within the parenthesis in Eq. (5) becomes zero and J_{1R} vanishes in this limit. In the opposite limit of high k_{2R}, $J_{1R} \to k_{1R}\tilde{P}_E$. However, as $k_{2R}k_{3R}$

(and δ_is) is fixed, k_{3R} is very low in this limit. Then, for a low catalytic rate constant of the wrong pathway, k_{4W}, one gets $\tilde{P}_E << 1$ (see Sec. 1 of the supplementary material) and J_{1R} again becomes very small. Similar arguments apply for J_{1W}. Hence, $J_{1R(W)}$ varies in a non-monotonic manner. Fluxes for the other steps are also expected to follow this trend at NESS [see Eq. (3)]. The effects of variations of the forces are insignificant compared to those of the corresponding fluxes due to the logarithmic dependence of the former on the rate constants (see the supplementary material). Hence, according to Eq. (2), the EPR should also exhibit a *non-monotonic* variation as exhibited by the fluxes. Following similar procedure, it can be shown that for fixed k_{3R}/k_{-1R} and δ_is, the fluxes become very small in the limits of both high and low k_{-1R}. Hence, the EPR is expected to show a non-monotonic variation in this case too. As the EPR is always non-negative, non-monotonic variation means that it should pass through a maximum. The catalytic rates, $v_{R(W)}$, and accuracy, being the ratio of such rates, can also show non-monotonic variations. For our general network far from equilibrium, the accuracy varies between two limits for the two cases of parameter variation described in the previous paragraph (see Sec. 3 of the supplementary materialfor details). However, the accuracy remains invariant when k_{-1R} is varied keeping the ratio k_{2R}/k_{-1R} and δ_i factors fixed (see Sec. 3 of the supplementary material). Hence, this case of parameter variation is not studied further.

III. DISSIPATION CONTROLS THE OPTIMIZATION OF CATALYTIC RATE IN A t-RNA SELECTION NETWORK

The feasibility of non-monotonic variation of EPR as well as the catalytic rate in KPR networks poses a very interesting question: *do both the quantities get optimized in similar parameter regimes?* To investigate this, we take a real biological proofreading network and experimental data of aminoacyl(aa)-tRNA selection during protein translation by *E. coli* wild-type ribosome.[23] The correspondence of this network to that depicted in Fig. 1 is as follows: E represents the ribosome with mRNA and R(W) substrate is the cognate (near-cognate) aa-tRNA in the ternary complex with elongation factor-Tu and GTP. GTP hydrolysis in step-2 forms the complex $ER^*(EW^*)$, which can take one of the two paths. Either it can form the product $P_R(P_W)$, i.e., the peptide chain elongated by one right (wrong) amino acid, or it can dissociate in the proofreading step (step-3), discarding the aa-tRNA. The rate constants based on available experimental data[23] are listed in the supplementary material (Table 1). We set the reverse rate constants of the strongly irreversible steps as $k_{-2R} = 10^{-4}$, $k_{-3R} = 10^{-3}$, $k_{-4R} = 10^{-5}$, all in s^{-1}, with $\delta_{-2} = 1$, $\delta_{-4} = 10^{-4}$. This automatically fixes $\delta_{-3} = 2.74 \times 10^{-3}$ as $\Pi_{i=1}^{3} \frac{\delta_i}{\delta_{-i}} = 1$.

The results of our numerical investigations are shown in Fig. 2. The EPR indeed passes through a maximum as a function of k_{2R} with fixed $k_{2R}k_{3R}$ and δ_is, as shown in Fig. 2(a). The catalytic rate of right product formation, v_R, also varies in a non-monotonic fashion. Remarkably, the system, shown by the red dot, *resides pretty close to the maximum of EPR as well as* v_R, depicted in Figs. 2(b) and 2(c), respectively. The accuracy, on the other hand, is far from the maximum-possible value. This feature corroborates with the recent finding that KPR networks prefer speed over accuracy.[21]

The proofreading step (step-3) resets the system without forming the desired product but already consuming ATP in the previous step (step-2). Although essential for the wrong pathway, the proofreading step is redundant for the right one. In contrast, catalysis (step-4) should be the priority for the right path and negligible for the wrong one. Therefore, it is important to study the contributions of the proofreading steps and the catalytic steps toward the EPR of the whole network. To this end, we introduce the function denoted

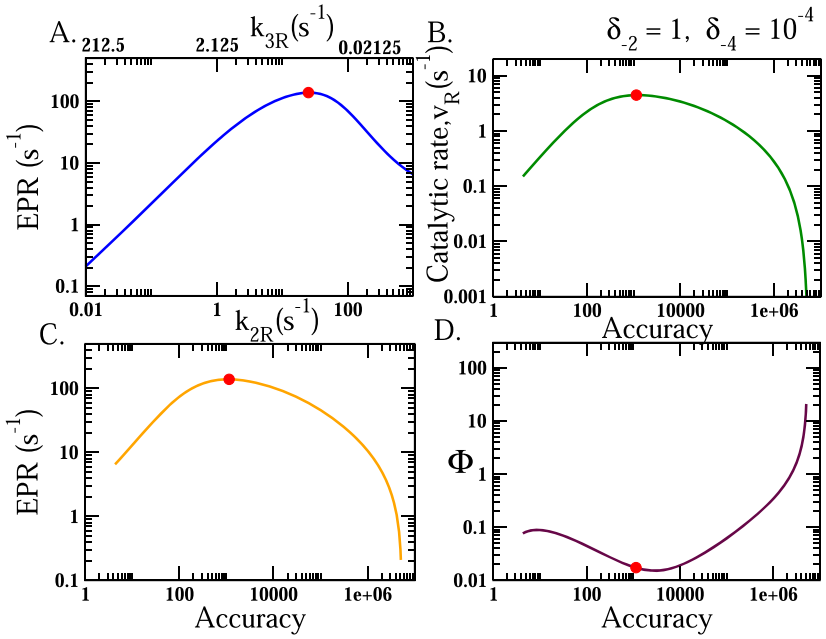

FIG. 2. (a) Variation of EPR as a function of k_{2R} with fixed $k_{2R}k_{3R}$ and δ_is. Variations of (b) catalytic rate, (c) EPR, and (d) dissipative partitioning, Φ, are depicted against accuracy. The red dot shows the position of the tRNA selection network.

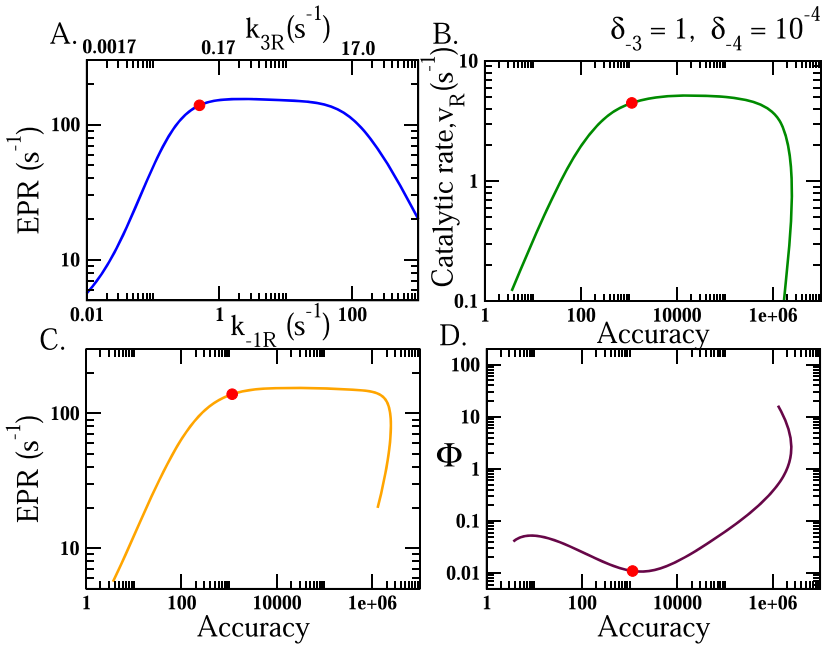

FIG. 3. (a) Variation of EPR as a function of k_{-1R} with fixed k_{3R}/k_{-1R} and δ_is. Variations of (b) the catalytic rate, (c) EPR, and (d) dissipative partitioning, Φ, are shown against accuracy. The red dot represents the tRNA selection network.

as dissipative partitioning, Φ,

$$\Phi = \left(\frac{\sigma_{3R} + \sigma_{3W}}{\sigma_{4R} + \sigma_{4W}}\right) = \left(\frac{J_{3R}\Delta\mu_{3R} + J_{3W}\Delta\mu_{3W}}{J_{4R}\Delta\mu_{4R} + J_{4W}\Delta\mu_{4W}}\right), \tag{6}$$

defined in light of the energetic cost of proofreading in terms of fluxes.[24] Here, $\sigma_{3R(3W)}$ is the EPR at NESS due to step-3 of the right (wrong) pathway; $\sigma_{4R(4W)}$ is the corresponding quantity for step-4. We plot the variations of Φ and accuracy in Fig. 2(d) as a function of k_{2R} with fixed $k_{2R}k_{3R}$ and δ_is. It is clear from Fig. 2(d) that the function Φ *not only passes through a local minimum but also the system (red dot) resides near it*. We have tested variations of all the quantities plotted in Fig. 2 in the same manner except for the change in the parameter δ_{-3} to $\delta_{-3} = 1$; then, $\delta_{-2} = 2.74 \times 10^{-3}$. All the features depicted in Fig. 2 remain essentially unchanged with this wide variation of free parameters (see Fig. 1 in the supplementary material).

Next, we explore the variations of all the quantities as a function of k_{-1R} keeping k_{3R}/k_{-1R} and δ_is fixed. One can clearly see from the results, shown in Fig. 3, that the system again lies very close to the maximum of the EPR and the catalytic rate but away from maximum accuracy [see Figs. 3(a)–3(c)]; it also lies near the minimum of Φ as given in Fig. 3(d). If the system moves further right along the curves in Figs. 3(b) and 3(c), it would lead to higher accuracy with similar EPR and rates. However, the constraint of minimization of Φ prevents this. We confirmed the robustness of the results shown in Figs. 2 and 3 by varying the free parameter δ_{-4} over the range 10^{-5}–1. The corresponding variations of EPR and Φ against accuracy are given in the supplementary material (Fig. 2). We further point out an interesting feature regarding the evolution of accuracy and Φ for both the cases of parameter variation mentioned above: they vary in a similar fashion for (i) fixed $k_{2R}k_{3R}$ and (ii) fixed k_{3R}/k_{-1R} with a fixed set of δ_is [see Figs. 2(d) and 3(d) for a comparison]. The theoretical justification of this behavior is given in the supplementary material. Similar behavior of Φ in both the cases of parameter variation shows the importance of this function during the evolution of rates of the various steps of the KPR network.

IV. CONCLUSION

In this work, we have theoretically analyzed a generalized version of the Hopfield KPR network under the biologically meaningful constraint of fixed chemical potential difference over the resetting cycles. In the chosen strategy of parameter variation, the EPR for the whole KPR network shows a non-monotonic variation and passes through a maximum as a function of rate constants of the various steps involved, viz., ATP hydrolysis, proofreading, and dissociation. The catalytic rate also goes through a maximum for similar variation of parameters. Considering the model parameters equal to those obtained from a real biological network of tRNA selection in wild-type *E. coli* ribosome revealed that it lies close to the maximum of both the EPR and the catalytic rate. The accuracy, on the other hand, is far from the maximum-possible limit. As a matter of original interest, proofreading was proposed to explain the high accuracy of various biological polymerization processes through non-equilibrium driving. However, our results show that a crucial advantage of being out-of-equilibrium is the dissipation-controlled optimization of the catalytic rate of protein synthesis. Remarkably, the real system not only tends to maximize the dissipation, i.e., the EPR, over the whole network to achieve the highest catalytic rate but also, simultaneously, minimizes the contribution of the proofreading step toward the dissipation. Therefore, our findings give useful insights into the categorical roles of (i) the distribution of the EPR among the various steps of the tRNA selection network to maximize the catalytic rate and (ii) a thermodynamic necessity of the preference shown

for speed over accuracy during protein synthesis. It will be important to test our results in more complex models of KPR, which, we believe, will provide further useful scenarios of the interplay of thermodynamic and kinetic features in such networks.

SUPPLEMENTARY MATERIAL

See the supplementary material for the (i) complete derivation of the formula of non-equilibrium steady state (NESS) probabilities, fluxes, and their variations; (ii) a detailed description of the variation of the catalytic rate for the right pathway; and (iii) variation of accuracy for different cases.

ACKNOWLEDGMENTS

K.B. acknowledges Professor Oleg A. Igoshin and Professor Anatoly B. Kolomeisky for useful discussions. The authors thank two anonymous reviewers for their useful comments and suggestions.

REFERENCES

[1]K. G. Denbigh, *The Principles of Chemical Equilibrium* (Cambridge University Press, 1981).
[2]S. R. De Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (Dover Publications, New York, 1984).
[3]L. Jiu-li, C. Van den Broeck, and G. Nicolis, Z. Phys. B: Condens. Matter **56**, 165 (1984).
[4]H. Qian, Annu. Rev. Phys. Chem. **58**, 113 (2006).
[5]H. Ge, M. Qian, and H. Qian, Phys. Rep. **510**, 87 (2012).
[6]H. Ge and H. Qian, Phys. Rev. E **81**, 051133 (2010).
[7]K. Banerjee, B. Das, and G. Gangopadhyay, J. Chem. Phys. **136**, 154502 (2012).
[8]B. Das, K. Banerjee, and G. Gangopadhyay, J. Chem. Phys. **139**, 244104 (2013).
[9]T. A. Kunkel, J. Biol. Chem. **279**, 16895 (2004).
[10]H. S. Zaher and R. Green, Cell **136**, 746 (2009).
[11]J. J. Hopfield, Proc. Natl. Acad. Sci. U. S. A. **71**, 4135 (1974).
[12]J. Ninio, Biochimie **57**, 587 (1975).
[13]A. Murugan, D. A. Huse, and S. Leibler, Proc. Natl. Acad. Sci. U. S. A. **109**, 12034 (2012).
[14]R. Rao and L. Peliti, J. Stat. Mech.: Theory Exp. **2015**, P06001.
[15]P. Sartori and S. Pigolotti, Phys. Rev. X **5**, 041039 (2015).
[16]H. Qian, J. Mol. Biol. **362**, 387 (2006).
[17]G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, Nat. Phys. **8**, 422 (2012).
[18]C. H. Bennett, BioSystems **11**, 85 (1979).
[19]F. Cady and H. Qian, Phys. Biol. **6**, 036011 (2009).
[20]P. Sartori and S. Pigolotti, Phys. Rev. Lett. **110**, 188101 (2013).
[21]K. Banerjee, A. B. Kolomeisky, and O. A. Igoshin, Proc. Natl. Acad. Sci. U. S. A. **114**, 5183 (2017).
[22]J. D. Mallory, A. B. Kolomeisky, and O. A. Igoshin, J. Phys. Chem. B **123**, 4718 (2019).
[23]H. S. Zaher and R. Green, Mol. Cell **39**, 110 (2010).
[24]M. A. Savageau and R. R. Freter, Proc. Natl. Acad. Sci. U. S. A. **76**, 4507 (1979).