Clustering and Optimal Arrangement of Enzymes in Reaction-Diffusion Systems

Alexander Buchner, Filipe Tostevin, and Ulrich Gerland*

Arnold Sommerfeld Center for Theoretical Physics and Center for Nanoscience, Ludwig-Maximilians-Universität München, Theresienstraße 37, D-80333 München, Germany

(Received 5 February 2013; published 16 May 2013)

Enzymes within biochemical pathways are often colocalized, yet the consequences of specific spatial enzyme arrangements remain poorly understood. We study the impact of enzyme arrangement on reaction efficiency within a reaction-diffusion model. The optimal arrangement transitions from a cluster to a distributed profile as a single parameter, which controls the probability of reaction versus diffusive loss of pathway intermediates, is varied. We introduce the concept of *enzyme exposure* to explain how this transition arises from the stochastic nature of molecular reactions and diffusion.

DOI: 10.1103/PhysRevLett.110.208104 PACS numbers: 87.15.R−, 87.10.−e, 87.18.Nq

To efficiently catalyze multistep biochemical reactions, sets of enzymes have evolved to function synergistically. Cells not only keep concerted control over the concentrations and activities of enzymes in the same pathway, but often also arrange them in self-assembled multienzyme complexes [1]. Apart from the large molecular machines (polymerases, ribosomes, spliceosomes), one of the best-studied natural multienzyme complexes is the cellulosome, a complex where up to 11 different enzymes are arranged on a noncatalytic scaffolding protein [2]. This complex is assembled extracellularly by anaerobic bacteria to efficiently break down cellulose, the most abundant organic material on the planet. Similarly, enzyme complexes are used for intracellular metabolism [3]. However, neither the precise consequences of putting enzymes together into complexes are well understood, nor the degree to which complex formation confers a functional advantage in each case [4–7].

It has long been thought that physical association between collaborating enzymes might increase the effective reaction flux, minimize the pool of unwanted intermediate products, allow coordinate regulation by a single effector, and reduce transient time scales [8,9]. However, while enzymatic activity has been studied for over a century, suitable techniques to characterize such effects quantitatively have become available only recently. On the one hand, single-molecule enzymology allows us to monitor [10] and manipulate [11] the activity of individual enzyme molecules. On the other hand, enzyme molecules can be positioned with nanometer precision in artificial systems using "single-molecule cut-and-paste" [12] on two-dimensional surfaces or along one-dimensional channels, and with DNA origami structures even in three dimensions [13,14]. These experimental developments call for a theoretical analysis of the effects of spatial proximity and arrangement of enzymes, to uncover the principles for the design and optimization of multienzyme systems. Such principles could be applied to bioengineer systems that control biochemical reactions at will, such as for the production of drugs or biofuels [15,16]. Related issues also arise in the context of signaling proteins [17]; however, the functional criteria for the optimization of signaling systems are likely different [18,19].

Here, we ask under which conditions it is beneficial to localize enzymes rather than distribute them. Furthermore, what is the optimal arrangement and how does it depend on the system parameters? We base this study on simple-reaction diffusion models, which permit rigorous quantitative analysis, and assume the steady-state reaction flux is the single critical system property. Interestingly, this already leads to rich physical behavior, with a sharp transition from a regime in which it is optimal to cluster downstream enzymes in the vicinity of upstream enzymes, to a regime in which an extended enzyme profile generates a higher reaction flux. This behavior, which we explain by analyzing the "enzyme exposure" of molecules diffusing in the system, is a result of the stochastic nature of the reactions and diffusion of single molecules.

Clustered enzymes.—That colocalizing enzymes within the same pathway might indeed improve the efficiency of converting a substrate S into a final product P can be seen by considering a two-step reaction, $S \xrightarrow{E_1} I \xrightarrow{E_2} P$, as a minimal model where production of P via an intermediate I is catalyzed by the enzymes E_1 and E_2. Let us consider an E_1 molecule (or a small cluster thereof) as a local source of I molecules and describe the local arrangement of E_2 enzymes relative to E_1 by the distribution $e(\mathbf{r})$, normalized such that $E_T = \int e(\mathbf{r}) d\mathbf{r}$ is the total number of E_2 molecules per E_1 center. To determine the efficiency of an enzyme arrangement $e(\mathbf{r})$, we need to describe the reaction-diffusion dynamics of the density $\rho(\mathbf{r}, t)$ of intermediates. We assume simple diffusion, with coefficient D, and standard Michaelis-Menten kinetics [20] for the enzymatic reactions, with catalytic rate k_{cat} and Michaelis constant K_M for E_2. In the low-density regime, where the reaction term becomes linear, we then have

$$\partial_t \rho(\mathbf{r}, t) = D \nabla^2 \rho(\mathbf{r}, t) - \kappa e(\mathbf{r}) \rho(\mathbf{r}, t), \qquad (1)$$

with $\kappa = k_{\rm cat}/K_M$ measuring the enzyme efficiency. Intermediates will either react to form product or will be lost, either directly to the extracellular space (for extracellular enzymes) or across the cell membrane. We can implement this possible loss via an absorbing boundary condition, $\rho(r = R, t) = 0$, on a sphere with radius R that may be taken to infinity. On the other hand, intermediates are constantly generated by E_1 at the origin, with an average flux that we denote by J_1, yielding the source boundary condition $-D(4\pi r^2 \partial_r \rho)_{r=0} = J_1$. In the resulting nonequilibrium steady-state $\rho(\mathbf{r})$, product is generated at the rate

$$J_2 = \kappa \int_{r<R} e(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r}. \tag{2}$$

Let us assume, for the moment, that enzyme E_2 is spread over a spherical shell with radius $r_0 < R$. We then find a total product flux of

$$J_2 = \frac{J_1}{1 + \frac{4\pi D R r_0}{E_T \kappa (R - r_0)}} \overset{R \gg r_0}{\longrightarrow} \frac{J_1}{1 + \frac{4\pi D r_0}{E_T \kappa}}. \tag{3}$$

This result indicates that reducing r_0—arranging the E_2 molecules close to the E_1 center—can dramatically increase the flux if loss of intermediate products is a concern. Whether this effect is biologically relevant crucially depends on the characteristic length scale $r_c = E_T \kappa / 4\pi D$, where J_2 begins to saturate. Enzyme efficiencies can be up to $\kappa \sim 10^8\ \mathrm{M^{-1}\,s^{-1}}$ (although superefficient enzymes can achieve $\kappa \sim 10^{10}\ \mathrm{M^{-1}\,s^{-1}}$ [21]), while biomolecular diffusion constants are typically larger than $D \sim 10\ \mu\mathrm{m^2\,s^{-1}}$, such that with $E_T \sim 10\, E_2$ molecules per E_1 center, r_c is at most of nanometer scale, comparable to the size of enzymes. Thus even our simplified model, which does not include interenzyme interactions such as direct channeling [22], suggests that in realistic biochemical settings, J_2 will be strongly dependent on the distance between enzymes down to the scale of their own size.

On a microscopic scale, the simple reaction-diffusion description we have used above will break down, since steric effects and the specific enzyme structure become important. Nevertheless, we can exploit the coarse-grained model to address more general questions on a mesoscopic scale. In particular, it is intriguing to ask whether colocalization is in fact the optimal enzyme arrangement, and whether the behavior will change qualitatively when the enzyme kinetics become nonlinear.

Clustered vs uniform arrangements.—Let us focus on the one-dimensional version of Eq. (1). This is not only a natural starting point for a theoretical study, but also relevant experimentally, e.g. for "molecular factories" in quasi-one-dimensional channels within future "lab-on-a-chip" devices. Specifically, we consider the one-dimensional steady-state $\rho(x)$ of a finite system, $x \in [0, L]$, with source/sink boundaries, $-D(\partial_x \rho)_{x=0} = J_1$ and $\rho(L) = 0$. We compare different E_2 enzyme distributions $e(x)$ with the same mean density $\bar{e} = L^{-1} \int_0^L e(x) dx = E_T/L$. The behavior of the system is determined by the dimensionless control parameter $\alpha = \kappa \bar{e} L^2 / D$, which measures the relative importance of reactions and diffusion in shaping $\rho(x)$. When $\alpha < 1$, the system is dominated by diffusion, as the typical reaction time scale $(\kappa \bar{e})^{-1}$ is longer than the typical diffusion time $\sim L^2/D$ to the absorbing boundary. Conversely, for large α, reactions are fast compared to diffusive escape. In the limit of $\alpha \to \infty$, J_2 approaches J_1 independent of the spatial arrangement of enzymes.

We first compare the reaction flux of clustered enzymes, $e_c(x) = \bar{e}\delta(x/L)$, and uniform enzymes, $e_u(x) = \bar{e}$. As shown in Fig. 1, the clustered configuration achieves a larger flux for $\alpha \lesssim 9$. Surprisingly, for larger α, the uniform configuration achieves a higher reaction flux. Thus it is not always preferable to simply localize enzymes where the concentration of intermediate is highest, which always occurs at $x = 0$. Rather, when reactions are fast compared to diffusion, the intermediates can be consumed more efficiently if E_2 is uniformly distributed throughout the system.

Enzyme exposure.—To examine the origin of this transition, we consider the fate of a single I molecule introduced into the system at $t = 0$. Whether it will have reacted by time T depends on the concentration of E_2 enzymes, $e(x(t))$, to which it has been exposed along its trajectory $x(t)$: the probability that it has not reacted is $\exp[-\kappa \int_0^T e(x(t)) dt]$. Therefore, the probability of escaping the system can be decomposed into the likelihood of particular trajectories through the system, and the probability of no reaction occurring along each trajectory. Indeed, the relative likelihoods of escape and reaction can be recaptured if, rather than assuming that I is consumed by the enzyme, we instead propagate a diffusive trajectory until it hits the absorbing boundary at time τ, and subsequently determine whether or not a reaction would have occurred based on the rescaled total enzyme exposure $E = D(L^2 \bar{e})^{-1} \int_0^\tau e(x(t)) dt$ and reaction probability $p_r(E) = 1 - \exp(-\alpha E)$.

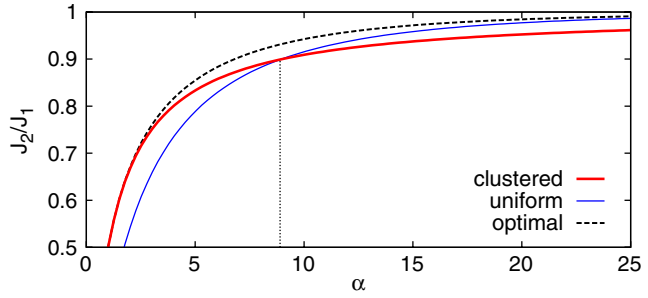

FIG. 1 (color online). Comparison of the reaction flux achieved by different enzyme profiles. A transition occurs at $\alpha \sim 9$ between regimes in which clustered and uniform enzyme profiles achieve a higher reaction flux. The optimal mixed enzyme distribution [Eq. (5) with $f = \alpha^{-1/2}$, dashed black line] achieves a still higher J_2 for intermediate values of α.

Given the stochasticity of diffusion, a given enzyme arrangement $e(x)$ will lead to a characteristic distribution of enzyme exposure, $P(E)$. For uniformly distributed enzymes, E is simply proportional to the time spent in the system, and $P(E)$ is therefore set by the distribution of escape times at the absorbing boundary $x = 1$ for a diffusing particle (see the Supplemental Material [23]),

$$P_u(E) = \sum_{n=0}^{\infty} \pi(-1)^n(2n+1)e^{-\pi^2(n+1/2)^2E}. \quad (4)$$

For a clustered configuration the appropriate distribution is found to be $P_c(E) = \exp(-E)$ (see the Supplemental Material [23]). Importantly, these distributions are independent of the reaction rate α, which enters into the reaction flux only via the reaction probability $p_r(E)$, which is in turn independent of the spatial arrangement of enzymes. Specifically, the reaction flux is given by $J_2 = J_1 \int_0^\infty P(E)p_r(E)dE$. Thus it is the interaction of these two distributions that determines which enzyme profile is preferable for a given value of α.

Figure 2 rationalizes the transition observed in Fig. 1. When $\alpha \ll 1$, such that $p_r(E \lesssim 1)$ is small, the majority of reaction events correspond to trajectories with large values of E. Compared to the uniform configuration, for which $P_u(E) \sim \exp(-\pi^2 E/4)$ for large E, the clustered configuration places more probability weight in the large-E tail of $P_c(E)$, and thus achieves a higher reaction flux when α is small. In the opposite limit of large $\alpha \gg 10$, only those

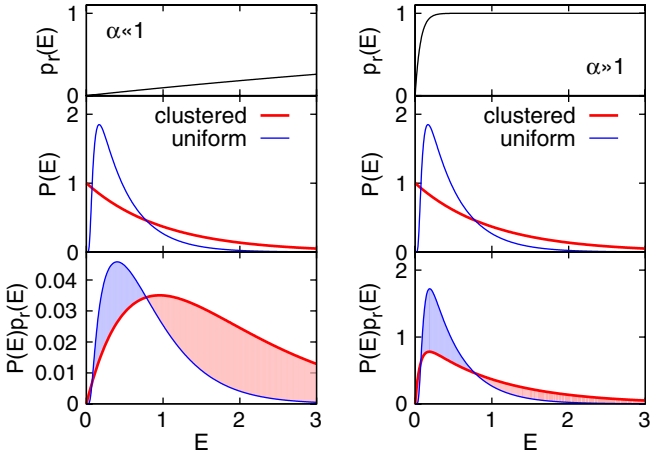

FIG. 2 (color online). Schematic depiction of the transition from a favorable clustered configuration ($\alpha \ll 1$, left) to the regime in which the uniform profile is preferable ($\alpha \gg 1$, right). Middle: when enzymes are clustered at $x = 0$ $P(E)$ has excess probability, compared to when enzymes are uniformly distributed, at small and large values of E. Bottom: the reaction flux is given by the integral of $P(E)p_r(E)$. For $\alpha \ll 1$ the extra probability in the large-E tail of $P(E)$ in the clustered configuration contributes more to J_2 than probability in the region $E < 1$. When $\alpha \gg 1$ only trajectories with $E \ll 1$ are subject to a low reaction probability, leading to a lower J_2 when enzymes are clustered.

trajectories with extremely small values of $E \ll 1$ have a significant probability of not reacting. Thus the uniform enzyme profile, for which $P_u(E \to 0) \to 0$, becomes preferable. The critical value of the transition, $\alpha \approx 9$, marks the point at which the reaction probability becomes large in the vicinity of the peak of $P_u(E)$.

Optimal profiles.—We have thus far compared only uniformly distributed and clustered configurations. However, it may be that another enzyme profile is able to achieve a reaction flux which is higher still. We therefore investigated what is the optimal enzyme distribution $e(x)$, for fixed $\bar{e}$, that maximizes the reaction flux J_2 (or alternatively, minimizes leakage $J_1 - J_2$). A direct analytic optimization of J_2 over $e(x)$ is not possible because of the nontrivial dependence of $\rho(x)$ on $e(x)$. We therefore studied the optimization of J_2 numerically on a discretized interval (see the Supplemental Material [23]).

These data show that for small $\alpha < 1$ the clustered configuration, with all enzymes colocalized with the source, is the optimal arrangement. Interestingly, the optimal profile undergoes a transition, distinct from that discussed above, at the critical value $\alpha = 1$. For $\alpha > 1$, in the optimal profile only a fraction of the available enzymes were clustered; the remaining enzymes were distributed approximately uniformly over an extended region with the enzyme density in this region equal to $\bar{e}$, as shown in Fig. 3.

Motivated by these numerical results, we studied enzyme profiles of the form

$$e(x) = \bar{e}\left\{f\delta\left[\frac{x}{L}\right] + \Theta\left[1 - f - \frac{x}{L}\right]\right\}, \quad (5)$$

where $\Theta(x)$ is the Heaviside function, and f is the fraction of enzymes that are clustered. We found that for this restricted class of profiles, the optimal profile indeed undergoes a transition from $f = 1$ for $\alpha \leq 1$ to $f = \alpha^{-1/2}$ for $\alpha > 1$. Examining the scaling of the fraction of enzymes that are clustered in the numerically optimized profiles, we find excellent agreement with this α scaling

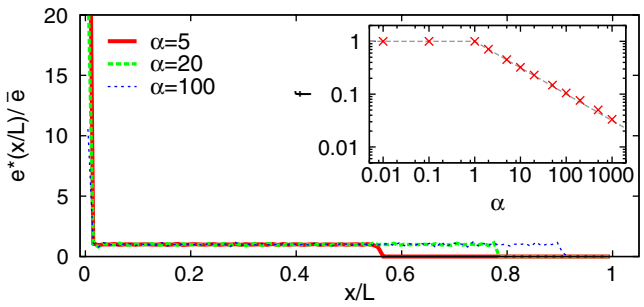

FIG. 3 (color online). Optimal enzyme density distribution for different values of α. Plotted profiles are the result of numerical optimization (see the Supplemental Material [23]) after 4×10^4 iterations with a lattice of 100 sites. Inset: the fraction of enzymes f located at the first lattice site in the numerically-optimized enzyme profile scales as $\alpha^{-1/2}$ for $\alpha > 1$.

(see Fig. 3 inset). The corresponding reaction flux tracks the envelope of the curves for the clustered and uniform configurations as α is varied (Fig. 1, dashed line).

The two distinct qualitative features of the optimal profile—the peak at $x = 0$ and the sharp decrease at $x = L(1 - \alpha^{-1/2})$—can be related to geometry of the system: enzymes cluster in the vicinity of the source and are excluded from the region nearest to the absorbing boundary. The distance from the end of the uniform enzyme domain to the boundary at $x = L$ scales with the typical diffusion length of substrate molecules in an enzyme density $\bar{e}$, which is $\sim L\alpha^{-1/2}$. If the enzyme concentration were to be uniform, $e(x) = \bar{e}$, substrate molecules that approach within this distance of the absorbing boundary have a high probability of diffusing out of the system rather than reacting. Any enzymes placed in this area contribute little to the reaction flux and can be used more effectively if relocated closer to the source.

We characterized $P(E)$ for mixed enzyme profiles of the form Eq. (5) by numerically sampling the enzyme exposure of continuous-time random walk trajectories on a lattice until their escape at $x = L$. The resulting distributions for different values of f are shown in Fig. 4. In the extreme cases of $f = 1$ and $f = 0$, the numerical results reproduce the analytic results of $P_c(E)$ and $P_u(E)$ above. At intermediate values of f, $P(E)$ retains a more pronounced large-E tail than $P_u(E)$, while still reducing the probability of extremely small E values relative to $P_c(E)$. As α is increased, the relative importance of these two features is reduced and increased, respectively. Thus the optimal $P(E)$ becomes more sharply peaked, corresponding to a smaller f.

So far we have considered only the case of linear reaction kinetics. In the nonlinear regime of the Michaelis-Menten kinetics, it is no longer possible to consider individual substrate trajectories independently since the reaction probability of a particular molecule depends on not only the local enzyme concentration but also the substrate density. Nevertheless, a qualitatively similar transition of the optimal enzyme distribution from clustered to distributed will occur provided the enzyme concentration

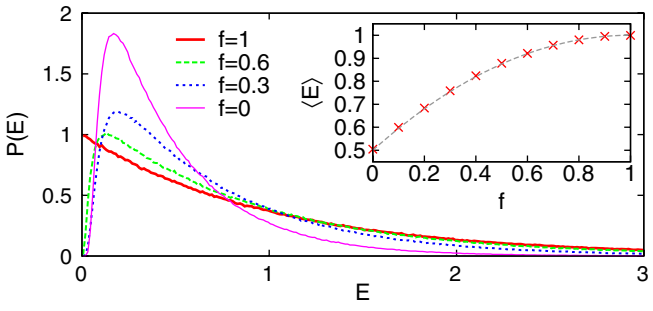

FIG. 4 (color online). Distributions $P(E)$ estimated from 2×10^6 simulated substrate trajectories subjected to an enzyme distribution $e(x_i) = \bar{e}\{fN\delta_{i,1} + \Theta[1 - f - (i/N)]\}$, with $N = 100$.

is not so low as to be saturated throughout the entire system, in which case the reaction current becomes independent of enzyme positioning.

Discussion.—In our model enzymatic pathway, the ultimate fate of each intermediate (I) molecule is either to react to product or to escape. For a given enzyme arrangement, the dimensionless parameter α controls the relative likelihood of these outcomes. Conversely, for each value of α there is an optimal enzyme arrangement that minimizes the loss of intermediates. In the small-α regime, where the reaction is slow and escape is likely, the best enzyme arrangement is a tightly clustered one. As α is increased, the system moves into the reaction-dominated regime and it becomes preferable to relocate some of the available E_2 enzymes away from the source. The transition of the optimal profile takes place at $\alpha \sim 1$. With a system size of $L \approx 100$ nm, α values in the range of 0.01–100 should be achievable in synthetic systems [12–14]. Thus it should be possible to directly test our results experimentally.

Intuitively, for large α, the more distant E_2 molecules may be interpreted as "backup enzymes" intended to catch the fraction of I molecules that were able to diffuse away from the cluster. The cost of removing some enzymes from the cluster is small since there remains a high probability of reaction for intermediates which spend a long time in the vicinity. Spreading these enzymes provides a larger benefit by recouping some of the escaped I molecules. Indeed, the optimal enzyme arrangement for $\alpha > 1$ is akin to a bet-hedging strategy, in the sense that the optimal placement for multiple E_2 molecules is not to cluster them all at the position where a single E_2 would do best, but instead hedge bets on the stochastic motion of their substrate by carefully distributing them.

Similar effects will also occur in systems with different geometries, including in higher dimensions. We have introduced the integrated "enzyme exposure" as a quantitative tool to characterize the effects of different enzyme arrangements. Importantly, the optimal enzyme profile does not necessarily maximize the average enzyme exposure (see Fig. 4 inset). Rather, it is the matching between the shape of the enzyme exposure distribution and the reaction probability that is key. While properties of diffusion such as recurrence change with dimension, the qualitative picture that clustered and distributed enzymes lead, respectively, to monotonically decaying and sharply peaked enzyme exposure distributions remains unchanged. Therefore the underlying physics of the transitions described is generic, although the magnitude of the effects will vary with the specific system. The concept of enzyme exposure provides a general framework for understanding the behavior of many different scenarios.

We have seen that the optimal enzyme distribution is determined by the distributions of timing of reaction and diffusion events. These are intrinsic single-molecule properties. Thus, we expect that the optimal enzyme profile

would remain unchanged if we considered instead discrete substrate and enzyme molecules. The only difference is that for finite numbers of enzyme molecules, $e(x)$ cannot be chosen arbitrarily but instead only certain discrete values are permitted. Thus $P(E)$ cannot be varied continuously, but rather one of a specific ensemble of allowed distributions must be chosen. While this will not change the qualitative behavior of the optimal profile as the system parameters are varied, it may quantitatively alter its shape for given parameter values. We leave this as a topic of future studies.

A. B. and F. T. contributed equally to this work. This research was supported by the German Excellence Initiative via the program "Nanosystems Initiative Munich" and the German Research Foundation via the SFB 1032 "Nanoagents for Spatiotemporal Control of Molecular and Cellular Reactions."

*gerland@lmu.de

[1] P. A. Srere. Annu. Rev. Biochem. **56**, 89 (1987).
[2] E. A. Bayer, H. Chanzy, R. Lamed, and Y. Shoham, Curr. Opin. Struct. Biol. **8**, 548 (1998).
[3] M. E. Campanella, H. Chu, and P. S. Low. Proc. Natl. Acad. Sci. U.S.A. **102**, 2402 (2005).
[4] A. Cornish-Bowden, Eur. J. Biochem. **195**, 103 (1991).
[5] P. Mendes, D. B. Kell, and H. V. Westerhoff, Eur. J. Biochem. **204**, 257 (1992).
[6] A. Cornish-Bowden and M. L. Cardenas, Eur. J. Biochem. **213**, 87 (1993).
[7] P. Mendes, D. B. Kell, and H. V. Westerhoff. Biochim. Biophys. Acta **1289**, 175 (1996).
[8] F. H. Gaertner, Trends Biochem. Sci. **3**, 63 (1978).
[9] R. Heinrich, S. Schuster, and H. Holzütter, Eur. J. Biochem. **201**, 1 (1991).
[10] S. Xie, Single Mol. **2**, 229 (2001).
[11] H. Gumpp, E. M. Puchner, J. L. Zimmermann, U. Gerland, H. E. Gaub, and K. Blank, Nano Lett. **9**, 3290 (2009).
[12] S. K. Kufer, E. M. Puchner, H. Gumpp, T. Liedl, and H. E. Gaub, Science **319**, 594 (2008).
[13] J. Müller and C. M. Niemeyer. Biochem. Biophys. Res. Commun. **377**, 62 (2008).
[14] J. Fu, M. Liu, Y. Liu, N. W. Woodbury, and H. Yan, J. Am. Chem. Soc. **134**, 5516 (2012).
[15] R. J. Conrado, J. D. Varner, and M. P. De Lisa, Curr. Opin. Biotechnol. **19**, 492 (2008).
[16] P. P. Peralta-Yahya, F. Zhang, S. B. del Cardayre, and J. D. Keasling, Nature (London) **488**, 320 (2012).
[17] D. Bray, Annu. Rev. Biophys. Biomol. Struct. **27**, 59 (1998).
[18] S. B. van Albada and P. R. ten Wolde, PLoS Comput. Biol. **3**, e195 (2007).
[19] A. Mugler, A. Gotway Bailey, K. Takahashi, and P. R. ten Wolde, Biophys. J. **102**, 1069 (2012).
[20] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics* (Wiley-Blackwell, Weinheim, Germany, 2012), 4th ed.
[21] M. E. Stroppolo, M. Falconi, A. M. Caccuri, and A. Desideri, Cell Mol. Life Sci. **58**, 1451 (2001).
[22] X. Huang, H. M. Holden, and F. M. Raushel, Annu. Rev. Biochem. **70**, 149 (2001).
[23] See Supplemental Material http://link.aps.org/supplemental/10.1103/PhysRevLett.110.208104 for details of the calculation of $P(E)$ and of the numerical optimization procedure.