Are DNA Transcription Factor Proteins Maxwellian Demons?

Longhua Hu,* Alexander Y. Grosberg,* and Robijn Bruinsma†
*Department of Physics, University of Minnesota, Minneapolis, Minnesota; and †Department of Physics and Astronomy, University of California, Los Angeles, California

ABSTRACT Transcription factor (TF) proteins rapidly locate unique target sites on long genomic DNA molecules—and bind to them—during gene regulation. The search mechanism is known to involve a combination of three-dimensional diffusion through the bulk of the cell and one-dimensional sliding diffusion along the DNA. It is believed that the surprisingly high target binding rates of TF proteins relies on conformational fluctuations of the protein between a mobile state that is insensitive to the DNA sequence and an immobile state that is sequence-sensitive. Since TFs are not able to consume free energy during their search to obtain DNA sequence information, the Second Law of Thermodynamics must impose a strict limit on the efficiency of passive search mechanisms. In this article, we use a simple model for the protein conformational fluctuations to obtain the shortest binding time consistent with thermodynamics. The binding time is minimized if the spectrum of conformational fluctuations that take place during the search is impedance-matched to the large-scale conformational change that takes place at the target site. For parameter values appropriate for bacterial TF, this minimum binding time is within an order-of-magnitude of a limiting binding time corresponding to an idealized protein with instant target recognition. Numerical estimates suggest that typical bacteria operate in this regime of optimized conformational fluctuations.

INTRODUCTION

The ability of bacteria to respond within minutes to changes in their environment relies on genetic switches that are controlled by transcription factors (TFs). TFs are proteins that—after activation by an environmental change—are able to locate a specific region (the operator sequence) along the bacterial genome and bind to it, thereby regulating the expression of a gene (or group of genes) adjacent to that region (1,2). The number of copies of a TF protein associated with a specific gene varies, but typically it is in the range of 10^2, corresponding to a concentration in the range of 0.1 μM. Because bacterial genomes have a size in the range of 10^7 sites, a TF must be able to scan the DNA for the target site at a rate of 10^5 sites per second or faster to at least one of them to reach the target site within seconds. Note that following the search for the target site, the TF still has to bind to the target site to regulate the expression of the gene.

A series of classical articles on the search process (2–4) culminated in the work of Berg et al. (5,6) and von Hippel and Berg (7) who showed—for the canonical case of the *lac* repressor protein of the bacterium *E. coli*—that the search process takes place not by straightforward three-dimensional diffusion to the target binding site but rather by a slide-jump combination of one-dimensional diffusional sliding along the DNA chain alternating with three-dimensional diffusional jumps between different DNA segments. By restricting part of the search to the one-dimensional target space, the binding rate is effectively enhanced with respect to a pure three-dimensional search, while the three-dimensional jumps reduce the repetitive visits to the same sites that characterize purely one-dimensional diffusive searches. This scenario is made possible by a modest, nonspecific electrostatic affinity between the TF and duplex DNA. More recently, one-dimensional diffusional sliding on DNA was directly observed in single molecule experiments (8–10). Berg et al. (5,6) and von Hippel and Berg (7) also provided evidence that, under physiological conditions, the search time has a minimum with respect to the strength of this nonspecific affinity, which may be the result of evolutionary optimization under selective pressure. Subsequent structural studies (11) have shown that the DNA-binding domains of the *lac* repressor are subject to strong conformational fluctuations when the protein is in contact with nonoperator DNA. If the binding domain is in contact with operator sequence DNA, then the protein can undergo a large-scale conformational change to a stable structure with direct contacts between the amino-acid side chains and the DNA bases.

It would seem obvious that the delay time between activation and binding of a TF to the operator sequence (i.e., binding time) is minimized by maximizing the one-dimensional diffusion constant D_1. However, simply increasing the transport rate will impair the accuracy, or fidelity, with which the protein can distinguish a right from a wrong site. Specifically, if the binding of a TF to the target site is characterized by a certain rate Ω, then the protein is likely to overshoot the target site if the jump rate D_1/a^2 between sites, with a the spacing between protein binding sites, is large compared to Ω. Similar conflicts between process speed and process fidelity are familiar from DNA duplication and transcription where increased reaction rates increase the number of duplication and transcription errors.

Submitted January 17, 2008, and accepted for publication April 7, 2008.

Address reprint requests to Alexander Y. Grosberg, Tel.: 612-624-7542; E-mail: grosberg@physics.umn.edu.

Editor: Arup Chakraborty.

0006-3495/08/08/1151/06 $2.00

doi: 10.1529/biophysj.108.129825

The search mechanism is known to involve a combination of three-dimensional diffusion through the bulk of the cell and one-dimensional sliding diffusion along the DNA. It is believed that the surprisingly high target binding rates of TF proteins relies on conformational fluctuations of the protein between a mobile state that is insensitive to the DNA sequence and an immobile state that is sequence-sensitive. Since TFs are not able to consume free energy during their search to obtain DNA sequence information, the Second Law of Thermodynamics must impose a strict limit on the efficiency of passive search mechanisms.

Slutsky and Mirny (12) proposed that conformational fluctuations of the protein could ease the conflict between speed and fidelity. If some conformations of the TF are sensitive to the DNA sequence while others are characterized by rapid transport, then the TFs might be able to scan the genome efficiently by appropriately flipping between the two types of conformations. The mechanism proposed by Slutsky and Mirny would be easy to envision for an active searcher, which spends free energy to gather information from the underlying DNA sequence and uses it to decide when it has to switch from the sliding mode to the recognition mode. However, TF proteins do not hydrolyze ATP or consume other forms of free energy during their search. It thus would seem that the Slutsky and Mirny mechanism requires TF proteins to act as Maxwellian Demons, able to gather information without expending free energy, but this is not permitted by the Second Law of Thermodynamics. The Second Law of Thermodynamics is, therefore, expected to impose a rigorous limit on the search efficiency of passive searchers. The aim of this article is to analyze how close this mechanism can approach limits of search efficiency imposed by fundamental principles of thermodynamics. We will address this question by examining a simple model for the conformational fluctuations, similar to that of Slutsky and Mirny (12), where the TF is allowed to adopt only two conformations (+ and −) when in contact with nonoperator DNA. Since the binding of TF to DNA involves a significant deformation of the double helix, the + and −states should be interpreted as states of a joint protein-DNA complex. For brevity, we will continue to refer to ''+'' and ''−'' as states of the protein. As illustrated in Fig. 1, in the +state, the protein is less ordered and only loosely associated with the DNA while it can slide along the DNA chain. In the −state, the protein is more ordered, closely associated with the DNA and immobile. A specific realization of a conformational fluctuation spectrum of this type was recently discussed for the *Ets*-DNA system (13). If the TF is in contact with the target operator sequence then, in addition to these two states, it also can undergo an irreversible conformational transition from the −state to the fully ordered final bound state. We will show that the shortest possible binding time in this model is controlled by a dimensionless binding rate $\omega \equiv \Omega ab/\sqrt{\pi K D_1 D_3}$, with D_3 the protein diffusion coefficient in bulk solution, D_1 the diffusion coefficient for one-dimensional transport along the DNA in

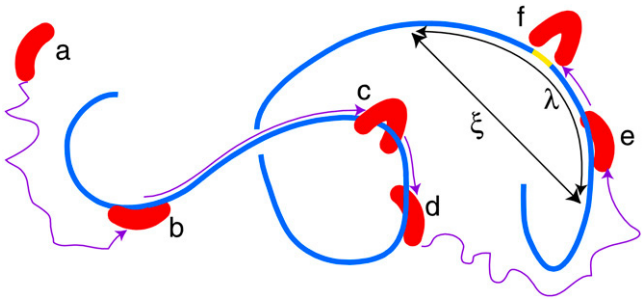

FIGURE 1 Schematic representation of the model. A protein moving diffusively through the cell volume (*a*) is adsorbed on genomic DNA (*b*) where it adopts one of two conformations: + and −. In the +conformation it is loosely associated with the DNA and can move by one-dimensional diffusion along the DNA chain (*b*) while in the −conformation (*c*) it is tightly associated with the DNA and is immobile. After returning to the +state, it restarts the sliding motion. The protein also can desorb from the chain (*d*) and return to three-dimensional diffusive motion. After a number of such cycles, the protein lands in the antenna region within a distance λ of the target binding site (*e*). After reaching the target site by one-dimensional diffusion it can undergo a large-scale irreversible conformational transition to the final bound state if it is in the −state (*f*).

the +state, K the equilibrium constant for the nonspecific protein-DNA interaction, and b the DNA-protein capture radius (more precisely, b is defined as the radius of a cylinder surrounding the DNA duplex such that a protein will be captured by the DNA, for example by electrostatic attraction, if the center of the protein is located inside the cylinder). If the dimensionless binding rate is comparable to one—or larger than one—then we can show that, for a particular value of the energy difference $\Delta E_{\pm}$ between the + and −conformations, the binding time can approach an absolute lower bound that corresponds to proteins having infinitely fast final binding rates. In other words, if the internal degrees of freedom of the protein in the sliding state are properly matched to the final binding rate, then the binding time of a TF can approach the shortest possible value allowed by thermodynamics provided the dimensionless binding rate is sufficiently large.

To demonstrate these claims, assume a cell of volume V containing a DNA genome of length L. The cell also contains a certain (low) concentration c of TF proteins that can bind reversibly and nonspecifically to the DNA. A protein whose center is located inside a cylindrical tube of radius b surrounding the duplex DNA will be assumed to be nonspecifically associated with the DNA. The fraction ϕ of the total cell volume occupied by the tube is of the order of Lb^2/V. There is also a single target site on the strand where the TF can bind irreversibly. We start by applying a fundamental theorem (14), which—in terms of our model—states that the mean waiting time for irreversible occupation of the target site, the quantity of interest to us, is equal to the inverse of a steady-state diffusion current of a different problem, namely one where the target site is replaced by a protein sink that constantly absorbs −state TF located at the target site at a rate Ω, while the protein concentration far from the target site is

maintained at a certain fixed value $c_3(\infty)$. The steady-state diffusion current, denoted by J_{3D}, into the target site for this second problem can be obtained from straightforward solution of the diffusion equation, which leads to the well-known Smoluchowski relation for the reaction rate of diffusion-limited chemical reactions:

$$J_{3D} \simeq 4\pi D_3 c_3(\infty)\xi. \quad (1)$$

Following Hu et al. (15), the effective target radius is defined as the radius of a sphere, surrounding the target site, that determines a crossover regime such that far outside the sphere adsorption of proteins onto the DNA chain is in equilibrium with evaporation of protein from the DNA chain while deep inside the sphere the absorption rate exceeds the evaporation rate. For the case of TF obeying slide-skip transport as in Berg et al. (5,6) and von Hippel and Berg (7), the size of this target sphere is determined by the condition that if a protein lands on a DNA segment inside the target sphere, after a three-dimensional diffusion step, then it typically reaches the target sink by pure one-dimensional diffusion where it gets absorbed before there is a chance for it to evaporate and leave the DNA. The length λ of DNA chain inside this target sphere—referred to as the antenna length—in general depends on the spatial organization of the genome. We will assume here the simple case of a straight genome, with ξ of $\sim\lambda$.

Our assumption that the antenna is straight is justified by the fact that the antenna length λ will turn out to be shorter than the DNA persistence length, which is close to 50 nm. If the antenna length would have exceeded the persistence length, then DNA conformation fluctuations would play an important role, and intersegmental transfer and Levi-flight transport would have to be considered. For a discussion of this more general case, see the literature (15–17).

Assuming the antenna is straight, its length has to be determined self-consistently but first we must establish a relation between $c_3(\infty)$ and the actual protein concentration c.

Far outside the target sphere the DNA-protein system is, by assumption, nearly in local thermal equilibrium, so one can determine the concentrations of adsorbed and free proteins purely from equilibrium considerations. If one views the association of the TF with DNA as a simple chemical reaction, then the concentration $\tilde{c}(\infty)$ of proteins adsorbed nonspecifically on the DNA and the concentration $c_3(\infty)$ of free proteins must be related to the reaction volume fraction ϕ by the Law of Mass Action for dilute chemical systems in thermodynamic equilibrium,

$$\frac{c_3(\infty)\phi}{\tilde{c}(\infty)} \simeq K, \quad (2)$$

with $\phi \ll 1$. The nonspecific protein-DNA equilibrium constant K depends strongly on the salt concentration (14), and other thermodynamic parameters, but it is independent of the protein and DNA concentrations. Since $c = c_3(\infty)+\tilde{c}(\infty)$, the concentrations of free and adsorbed proteins are now determined but it will be useful to replace the bulk concentration $\tilde{c}(\infty)$ of adsorbed proteins by the one-dimensional concentration $c_1(\infty) \simeq b^2\tilde{c}(\infty)/\phi$, the number of adsorbed proteins per unit length of DNA far from the target site. Solving for $c_1(\infty)$ and $c_3(\infty)$ gives $c_1(\infty) \simeq cb^2/K(1+\phi/K)$ and $c_3(\infty) \simeq c/(1+\phi/K)$, still for $\phi \ll 1$.

Deep inside the target sphere, the system is not in thermal equilibrium, with the adsorption rate of proteins from the bulk solution to the DNA exceeding the evaporation rate. The difference is matched by a one-dimensional diffusion current J_{1D} along the DNA chain toward the target site. To estimate this one-dimensional diffusional transport, note that if the interconversion rate between the + and −states is sufficiently rapid, then their respective occupancies can be approximated by the equilibrium Boltzmann distribution. The effective one-dimensional diffusion constant for transport along the chain—which we will denote by $\tilde{D}_1$—is then proportional to the Boltzmann probability $p(+)$ to find the protein in the +state. If $\mu \equiv \exp(-\Delta E_{\pm}/k_BT)$, then $p(+) = \mu/(1+\mu)$ and $\tilde{D}_1 \simeq D_1\mu/(1+\mu)$. Similarly, the effective target-site binding rate $\tilde{\Omega}$ is, under these same conditions, proportional to the probability $p(-) = 1 - p(+)$ to find the protein in the −state and $\tilde{\Omega} \simeq \Omega/(1+\mu)$.

Let $c_1(0)$ be the one-dimensional concentration at the target site. If the final binding rate were infinitely fast, then $c_1(0)$ would be zero but, because of the overshoot effect, this is no longer the case. If we view the surface of the target sphere as a matching region between the asymptotic regions far from the sink where the one-dimensional concentration approaches $c_1(\infty)$ and the region deep inside the target sphere near the sink where the one-dimensional concentration approaches $c_1(0)$, then we can estimate the one-dimensional concentration gradient as $[c_1(\infty) - c_1(0)]/\lambda$. It follows that the one-dimensional diffusion current toward the sink equals

$$J_{1D} \simeq \tilde{D}_1\frac{c_1(\infty) - c_1(0)}{\lambda}. \quad (3)$$

The number of proteins absorbed per second by the sink itself, J_s, is $\sim ac_1(0)\tilde{\Omega}$, with a the spacing between protein binding sites. Conservation of the number of proteins requires the three currents J_{3D}, J_{1D}, and J_s to be equal to each other (6), so

$$J_{3D} = J_{1D} = J_s. \quad (4)$$

Equating the one-dimensional diffusion current with the sink current allows us to eliminate $c_1(0)$ with the result:

$$J_{1D} \simeq \frac{\tilde{D}_1 c_1(\infty)}{\lambda}\left[\frac{\tilde{\Omega}}{\tilde{\Omega} + \tilde{D}_1/a\lambda}\right]. \quad (5)$$

The factor in front of the square brackets is the diffusion current in the absence of overshoot. The importance of overshoot is thus determined by the dimensionless number $a\lambda\tilde{\Omega}/\tilde{D}_1$. Since $\lambda^2/\tilde{D}_1$ is the typical time spent by a protein

diffusing along the antenna, it follows that $a\lambda/\tilde{D}_1$ is the typical time spent near the target site so $a\lambda\tilde{\Omega}/\tilde{D}_1$ is the product of the typical time spent near the target site with the effective absorption rate. The term inside the square brackets can then be understood as the probability for a protein in the antenna region to be trapped by the target.

We have been assuming that the energy difference between + and −states, or μ, does not depend on the DNA sequence. The effects of sequence-dependent, one-dimensional transport are discussed in the literature (18,19).

Equating the one-dimensional and three-dimensional currents provides us with a self-consistency condition that determines both the size of the antenna length λ and the reaction rate. Solving for λ using Eqs. 1 and 5 and using $c_1(\infty)/c_3(\infty) \simeq b^2/K$ gives the antenna length:

$$\lambda = \sqrt{\frac{b^2}{K}\frac{\tilde{D}_1}{4\pi D_3} + \left(\frac{\tilde{D}_1}{2\tilde{\Omega}a}\right)^2} - \frac{\tilde{D}_1}{2\tilde{\Omega}a}. \qquad (6)$$

The maximum value, $\lambda_\infty = \sqrt{b^2 D_1/4\pi K D_3}$, is reached for infinite Ω and infinite μ.

It will be helpful to express the binding rate $J_{3D} \simeq 4\pi D_3 c_3(\infty)\xi$ in dimensionless units as $A \equiv J_{3D}/(4\pi D_3 ca)$, with $4\pi D_3 ca$ the Smoluchowski limiting rate of a conventional three-dimensional diffusive search for an absorber target of radius a (the spacing between binding sites), so A can be viewed as a reaction amplification or enhancement factor. This enhancement factor can be expressed as a simple function of the dimensionless binding rate $\omega = \Omega ab/\sqrt{\pi K D_1 D_3}$ and the Boltzmann factor $\mu = \exp(-\Delta E_\pm/k_B T)$:

$$A(\omega,\mu) \simeq A_\infty \frac{1}{\omega}\left(\sqrt{\omega^2\frac{\mu}{1+\mu} + \mu^2} - \mu\right). \qquad (7)$$

Here $A_\infty = (b/a)\sqrt{KD_1/4\pi D_3}/(K+\phi)$ is the maximum value of the enhancement factor, corresponding to $\lambda = \lambda_\infty$ with both μ and Ω infinite. We will examine the amplification factor $A(\omega, \mu)$ as a function of the nonspecific equilibrium constant K and the occupation ratio μ of the + state and −state, rather than ω and μ, because these are physical parameters characterizing the interaction between the TF and the DNA that are expected to be sensitive to specific point mutations of the TF amino-acid sequence through their exponential dependence on binding and activation energies. The contour lines of constant A as a function of K and μ in Fig. 2 show that there is a single, rather shallow maximum. This maximum is also seen in the lower inset of the same figure, where A is plotted against μ at constant optimal K. The physical origin behind the maximum of A with respect to K is, as discussed earlier, the fact that a combination of one-dimensional and three-dimensional diffusion minimizes the search time. By contrast, the maximum of A as a function of μ at $\mu_{opt} = (\sqrt{1+2\omega} - 1)/2$ is surprising because it might have been expected that for sufficiently long DNA, location of the target site always should be the rate-limiting step, in which case the optimal choice for μ would be infinite since that maximizes the effective one-dimensional diffusion constant $\tilde{D}_1 = D_1\mu/(1+\mu)$. It can be shown that the maximum

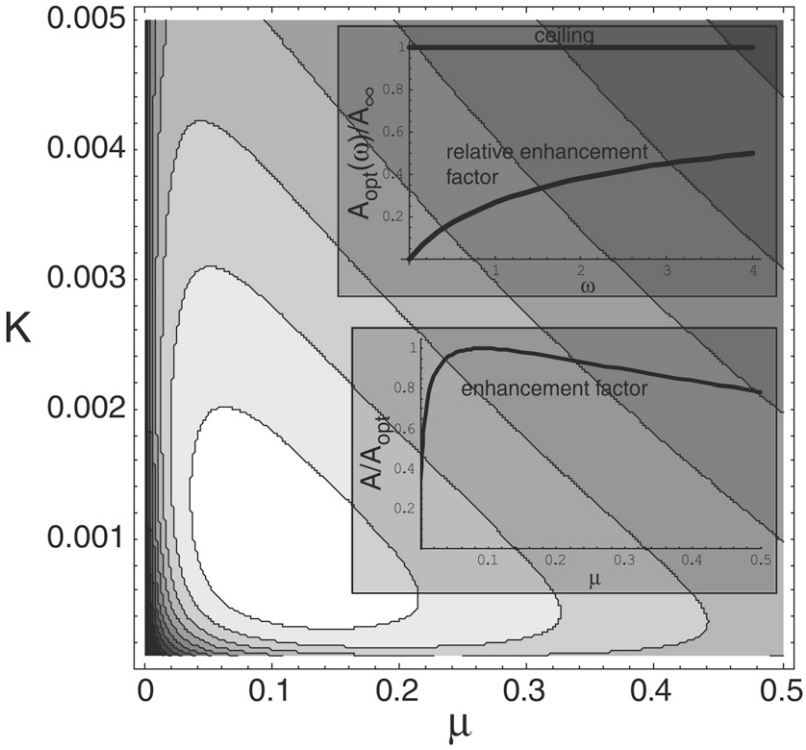

FIGURE 2 Contour plots of the transport enhancement factor A as a function of the equilibrium constant K and the occupation probability ratio μ of the + over –states for $D_1 = 10^{-9}$ cm^2/s, $D_3 = 3 \times 10^{-7}$ cm^2/s, $a = 0.34$ nm, $b = 5$ nm, $\phi = 0.01$, and $\Omega = 3 \times 10^3$ Hz. There is a shallow maximum at $\sim\mu = 0.1$ and $K = 10^{-3}$. The ratio of the transport enhancement factor at this maximum, A_{opt}, and the thermodynamic limiting enhancement factor, A_∞, equals 0.193. (*Upper inset*) Dependence of the ratio A_{opt}/A_∞ on the dimensionless binding rate ω. (*Lower inset*) Enhancement factor A against μ at the value of K corresponding to the maximum of the main figure.

with respect to μ actually is a form of impedance matching with the effective resistance of the one-dimensional diffusion search matched with the effective resistance of the binding process.

If μ adopts the optimal value $\mu_{\rm opt} = (\sqrt{1+2\omega} - 1)/2$, then the ratio $A_{\rm opt}/A_\infty$ of the optimal rate amplification factor and its maximum value is a function only of the dimensionless rate ω:

$$\frac{A_{\rm opt}(\omega)}{A_\infty} = 1 - \frac{1}{\omega}(\sqrt{1+2\omega} - 1). \qquad (8)$$

We thus are finding that the optimized amplification factor only depends on the system parameters in the form of the dimensionless combination $\omega = \Omega ab/\sqrt{\pi K D_1 D_3}$. The physical meaning of this key quantity becomes more transparent if we write it in the form $\omega = \Omega\tau$, where $\tau = ab/\sqrt{\pi K D_1 D_3}$ has the dimension of time. This timescale can be interpreted in the following way. The escape of a TF from DNA requires both release of the nonspecific bond between the protein and the DNA and three-dimensional diffusion out of the capture zone surrounding the DNA strand. If D_3/b^2 is the attempt rate for the escape, then KD_3/b^2 is the corresponding Arrhenius escape rate, with K the dissociation equilibrium constant. We thus can view b^2/KD_3 as the duration of an uninterrupted one-dimensional search tour. During each tour, the protein covers a distance of $\sim\sqrt{D_1 b^2/KD_3}$ by one-dimensional diffusion. Therefore, the time a protein spends on a given DNA site, including the target site, is $\sim ab/\sqrt{KD_1D_3}$, which is just our timescale τ. It follows that, if $\omega\tau$ is large compared to one, then the time spent by the protein on the target site is sufficiently long to practically guarantee capture. On the other hand, if $\omega\tau$ is small compared to one, then the protein is likely to overshoot the target site, and has to make many passes across the target before capture takes place. We thus indeed should expect a significant change in the capture efficiency when $\omega\tau$ is $\sim$1.

The dependence of $A_{\rm opt}/A_\infty$ on ω is shown in the upper inset of Fig. 2: $A_{\rm opt}$ is of the same order of magnitude as the theoretical limit A_∞ already for modest values of ω. This demonstrates our central claim: it is possible for the overall binding rate of a TF to approach the theoretical limiting value but only by a suitable choice of μ, and only if the dimensionless binding rate ω is of $\sim$1, or $>$1.

Are these two conditions realistic for typical TF? Typical values for the diffusion constants of bacterial TF are, according to the literature (8–10), $D_1 \approx 10^{-9}$ cm^2/s and $D_3 \approx 3 \times 10^{-7}$ cm^2/s. We can estimate the protein-DNA reaction volume fraction ϕ for *E. coli* by assuming it to be comparable to the DNA volume fraction ($\sim$1% within cell volume of the order of 1 cubic micrometer). The equilibrium constant can then be determined from the relation $c_3(\infty) \simeq c/(1+\phi/K)$ and the fact that it is known that $\sim$10% of the *lac* repressor proteins of *E. coli* are in solution (20), which means that K must be $\sim 10^{-3}$. If we assume a to be equal to the basepair spacing 0.34 nm, and estimate b as 5 nm, then the dimensionless binding rate ω is $\sim 10^{-4}\Omega$ with the binding rate Ω expressed in Hz. A large-scale protein conformational change typically involves millisecond-to-microsecond timescales, from which it follows that ω must lie in the range of 0.1–100. Note, from Fig. 2 that the optimal value for K is close to 10^{-3} for Ω in the kHz range. We conclude that the second condition can be satisfied under typical conditions. Next, the optimal occupation ratio $\mu_{\rm opt} = (\sqrt{1+2\omega} - 1)/2$ is in the range of 0.1–10 for ω in the range of 0.1–100. The corresponding optimal energy difference $\Delta E_\pm$ between the + and –states is then in the range of a few $k_{\rm B}T$, with $\Delta E_\pm$ positive for $\omega < 4$ but negative for $\omega > 4$. In either case, the structure of optimized TF bound to nonoperator DNA should be subject to strong thermal fluctuations. As we saw, this is indeed the case of the *lac* repressor (11), while a recent modeling study of the *Ets*-DNA system arrives at the same conclusion (13). The first condition can thus be satisfied as well under reasonable conditions. Finally, the measured *lac* repressor binding rates (5–7) are comparable to the limiting rate imposed by thermodynamics. We conclude that, under reasonable conditions, the binding rate of TF proteins can be of the same order of magnitude as the thermodynamically imposed limiting rate if the energy spectrum of conformational fluctuations is determined, under selective pressure, by minimization of the overall binding time.

APPENDIX: OPTIMAL SEARCHES ON COILED GENOMES

The above results were derived assuming that the antenna was straight. The role of coiled DNA conformation in the slide-jump search process was addressed systematically, for a variety of DNA conformations, in Berg et al. (6). The goal of this Appendix is to see how the optimization with respect to μ is implemented for coiled DNA conformations and, simultaneously, to demonstrate that the assumption of a straight antenna is justified in most physiological conditions.

The starting point is to consider the relation between antenna size, ξ, and antenna length along DNA, λ. For a straight antenna, $\xi = \lambda$, while for a Gaussian coiled antenna, $\xi \simeq \sqrt{p\lambda}$, where p is effective Kuhn segment of DNA, usually $p \approx 100$ nm. In the general case, we can write $\xi \simeq \ell(\lambda/\ell)^\nu$, with arbitrary power ν, $0 < \nu < 1$, and with suitably defined characteristic length ℓ. We note in passing that ℓ is not necessarily equal to the Kuhn segment p for instance, for the wormlike chain with excluded volume $\nu \approx 3/5$ and $\ell \approx \sqrt{pd}$, where d is the chain thickness.

In the arbitrary ν-case, it turns out that there is still a single dimensionless criteria, which generalizes ω, and which can be written in the form $\omega_\nu = \omega(b^2 D_1/\ell^2 4\pi K D_3)^{(1-\nu)/(2(1+\nu))}$. Using as before the equations $J_{\rm 3D} = J_{\rm 1D} = J_{\rm s}$ (4), but with a more general relation between ξ and λ, one can obtain the following equation for the normalized enhancement factor $\alpha = A/A_\infty$:

$$\alpha^{\frac{1+\nu}{\nu}} + \frac{2\mu}{\omega_\nu}\alpha = \frac{\mu}{1+\mu}. \qquad (9)$$

In the old case of a straight antenna, $\nu = 1$, this equation is quadratic and its solution is Eq. 7. In the general case this equation does not allow explicit solution, but we can still address optimization with respect to μ. It can be shown in a few lines of algebra that the optimized value of $\alpha = \alpha_{\rm opt} \equiv A_{\rm opt}/A_\infty$ satisfies the equation

$$\alpha_{\rm opt}^{\frac{1+\nu}{2\nu}} + \sqrt{\frac{2\alpha_{\rm opt}}{\omega_\nu}} = 1, \qquad (10)$$

which reproduces Eq. 8 if $\nu = 1$, and which in general yields the following asymptotics:

$$\alpha_{\rm opt} \simeq \begin{cases} 1 - \dfrac{2\nu}{1+\nu}\sqrt{\dfrac{2}{\omega_\nu}} & \text{when} \quad \omega_\nu \gg 1 \\ \dfrac{\omega_\nu}{2}\left[1 - 2\left(\dfrac{\omega_\nu}{2}\right)^{\frac{1+\nu}{2\nu}}\right] & \text{when} \quad \omega_\nu \ll 1 \end{cases}. \qquad (11)$$

Qualitatively, this is essentially the same behavior as illustrated in the upper inset of Fig. 2.

In the light of these results, we can now provide a more specific justification for the assumption of straight antenna which we used in the main text. For instance, for the in vitro situation of DNA in solution, DNA is straight ($\nu = 1$) at length scales up to about Kuhn segment $p \approx 100$ nm and becomes a Gaussian coil ($\nu = 1/2$) at larger length scales:

$$\xi \simeq \begin{cases} \lambda, \quad \text{or} \quad \nu = 1 & \text{when} \quad \lambda < p \\ \sqrt{\lambda p}, \quad \text{or} \quad \nu = 1/2,\ \ell = p & \text{when} \quad \lambda > p \end{cases}. \qquad (12)$$

Therefore, once we analyzed the straight antenna ($\nu = 1$) case and optimized the enhancement factor with respect to μ, we should check that the optimal antenna length is shorter than p. Omitting algebra, this yields the condition $(1 + \omega - \sqrt{1+2\omega})/\omega < (p/b)\sqrt{4\pi K D_3/D_1}$. Similarly, once we have optimized the enhancement for $\nu = 1/2$, we should check that the optimal antenna length is longer than p. This condition upon some algebra leads to exactly the opposite inequality $(1 + \omega - \sqrt{1+2\omega})/\omega > (p/b)\sqrt{4\pi K D_3/D_1}$. Thus, there is a smooth crossover between these two regimes. Given that the left-hand side of these two inequalities is a monotonically growing function of ω which approaches unity at very large ω, we arrive at the conclusion that the optimal antenna might become a Gaussian coil if and only if $(p/b)\sqrt{4\pi K D_3/D_1} < 1$. It is not difficult to see the physical meaning of this inequality: the average number of desorption events while diffusing one Kuhn length should be less than unity. Taking the typical numbers ($p = 100$ nm, $b = 5$ nm, $D_3 = 3 \times 10^{-7}$ cm^2/s, $D_1 = 10^{-9}$ cm^2/s), we see that $(p/b)\sqrt{4\pi K D_3/D_1} \approx 10^3\sqrt{K}$, which means the optimal antenna can become Gaussian only when $K \lesssim 10^{-6}$. Since values of K below this threshold are unlikely, the optimal antenna—no matter how big is ω—is typically shorter than the Kuhn segment, thus justifying our assumption in the main text that the antenna is more or less straight.

We thank Leonid Mirny for useful discussions and the Kavli Institute of Theoretical Physics and the Aspen Center for Physics, where this study was initiated, for their hospitality.

L.H. and A.G. acknowledge support by the Materials Research Science and Engineering Center Program of the National Science Foundation under award No. DMR-0212302 and R.B. acknowledge support by the National Science Foundation under DMR grant No. 0404507.

REFERENCES

1. Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 1994. Molecular Biology of the Cell. Garland, New York.
2. Adam, G., and M. Delbrück. 1968. Structural Chemistry and Molecular Biology. A. Rich and N. Davidson, editors. Freeman, New York.
3. Riggs, A. D., S. Bourgeois, and M. Cohn. 1970. The *lac* repressor-operator interaction: III. kinetic studies. *J. Mol. Biol.* 53:401–417.
4. Richter, P. H., and M. Eigen. 1974. Diffusion controlled reaction rates in spheroidal geometry: application to repressor-operator association and membrane bound enzymes. *Biophys. Chem.* 2:255–263.
5. Berg, O. G., R. B. Winter, and P. H. von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry.* 20:6929–6948.
6. Berg, O. G., R. B. Winter, and P. H. von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The *Escherichia coli* lac repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry.* 20:6961–6977.
7. von Hippel, P. H., and O. G. Berg. 1989. Facilitated target location in biological systems. *J. Biol. Chem.* 264:675–678.
8. Sokolov, I. M., R. Metzler, K. Pant, and M. C. Williams. 2005. Target search of *N* sliding proteins on a DNA. *Biophys. J.* 89:895–902.
9. Wang, Y. M., R. H. Austin, and E. C. Cox. 2006. Single molecule measurements of repressor protein 1D diffusion on DNA. *Phys. Rev. Lett.* 97:048302.
10. Elf, J., G. Li, and X. S. Xie. 2007. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science.* 316: 1191–1194.
11. Kalodimos, C. G., N. Biris, A. Bonvin, M. Levandoski, M. Guennuegues, R. Boelens, and R. Kaptein. 2004. Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science.* 305:386–389.
12. Slutsky, M., and L. A. Mirny. 2004. Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophys. J.* 87:4021–4035.
13. Levy, Y., J. N. Onuchic, and P. G. Wolynes. 2007. Fly-casting in protein-DNA binding: frustration between protein folding and electrostatics facilitates target recognition. *J. Am. Chem. Soc.* 129:738–739.
14. Pontryagin, L., A. Andronov, and A. Vitt. 1933. *Zh. Eksp. Teor. Fiz.* 165:3. Translated and reprinted in Noise in Nonlinear Dynamical Systems, Vol. 1, 1989. F. Moss and P. V. E. McClintock, editors. Cambridge University Press, Cambridge, UK.
15. Hu, T., A. Y. Grosberg, and B. I. Shklovskii. 2006. How proteins search for their specific sites on DNA: the role of DNA conformation. *Biophys. J.* 90:2731–2744.
16. Lomholt, M. A., T. Ambjornsson, and R. Metzler. 2005. Optimal target search on a fast-folding polymer chain with volume exchange. *Phys. Rev. Lett.* 95:260603.
17. Hu, T., and B. I. Shklovskii. 2007. How a protein searches for its specific site on DNA: the role of intersegment transfer. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 76:051909.
18. Slutsky, M., M. Kardar, and L. A. Mirny. 2004. Diffusion in correlated random potentials, with applications to DNA. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 69:061903.
19. Hu, T., and B. I. Shklovskii. 2006. How does a protein search for the specific site on DNA: the role of disorder. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 74:021903.
20. Kao-Huang, Y., A. Revzin, A. Butler, P. O'Conner, D. Noble, and P. H. von Hippel. 1977. Nonspecific DNA binding of genome-regulating proteins as a biological control mechanism: measurement of DNA-bound *Escherichia coli* lac repressor in vivo. *Proc. Natl. Acad. Sci. USA.* 74:4228–4232.