15.3: Search Times in Facilitated Diffusion

Consider a series of repetitive 1D and 3D diffusion cycles.[1] The search time for a protein to find its target is

$$t_s = \sum_{i=1}^{k} (\tau_{1D,i} + \tau_{3D,i})$$

where k is the number of cycles. If the genome has a length of M bases and the average number of bases scanned per cycle is n, the average number of cycles $\bar{k}$ = M/$\bar{n}$, and the average search time can be written as

$$\bar{t}_s = \frac{M}{\bar{n}} (\bar{\tau}_{1D,i} + \bar{\tau}_{3D,i}) \tag{15.3.1}$$

$\bar{\tau}$ is the mean search time during one cycle. If we assume that sliding occurs through normal 1D diffusion, then we expect that $\bar{n} \propto \sqrt{D_{1D}\bar{\tau}_{1D}}$, where the diffusion constant is expressed in units of bp^2/s. More accurately, it is found that if you executed a random walk with an exponentially weighted distribution of search times:

$$P(\tau_{1D}) = \bar{\tau}_{1D}^{-1} \exp(-\tau_{1D}/\bar{\tau}_{1D})$$
$$\bar{n} = \sqrt{4 D_{1D} \bar{\tau}_{1D}}$$
$$\bar{t}_s = \frac{M}{\sqrt{4 D_{1D} \bar{\tau}_{1D}}} (\bar{\tau}_{1D} + \bar{\tau}_{3D}$$

Let's calculate the optimal search time, t_{opt}. In the limits that $\bar{\tau}_1$ or $\bar{\tau}_3 \to 0$, you just have pure 1D or 3D diffusion, but this leads to suboptimal search times because a decrease in $\bar{\tau}_{1D}$ or $\bar{\tau}_{3D}$ leads to an increase in the other. To find the minimum search time we solve:

$$\frac{\partial \bar{t}_s}{\partial \tau_{1D}} = 0$$

and find that t_{opt} corresponds to the condition

$$\bar{\tau}_{1D} = \bar{\tau}_{3D}$$

Using this in eq. (15.3.1) we have

$$t_{opt} = \frac{2M}{\bar{n}} \bar{\tau}_{3D} = M \sqrt{\frac{\bar{\tau}_{3D}}{D_{1D}}}$$
$$\bar{n}_{opt} = \sqrt{4 D_{1D} \bar{\tau}_{3D}}$$

Now let's find out how much this 1D + 3D search process speeds up over the pure 1D or 3D search.

- **3D only:** $\bar{\tau}_{1D} \to 0 \qquad \therefore \bar{n} \to \sim 1$ leading to

$$\bar{t}_{3D} = M \bar{\tau}_{3D}$$

  Facilitated diffusion speeds up the search relative to pure 3D diffusion by a factor proportional to the average number of bases searched during the 1D sliding.

$$\frac{\bar{t}_{3D}}{(\bar{t}_s)_{opt}} = \frac{\bar{n}}{2}$$

- **1D only:** $\bar{\tau}_{3D} \to 0 \qquad \therefore \bar{n} \to M$, and

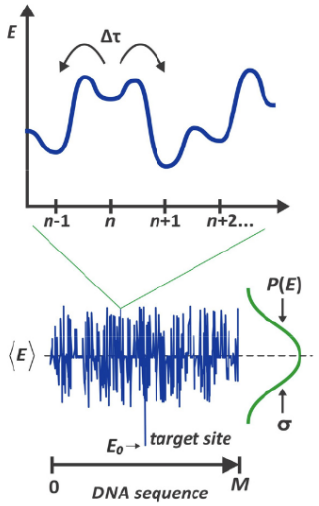
$$\bar{t}_1 D \approx \frac{M^2}{4 D_{1D}}$$
$$\frac{\bar{\tau}_{1D}}{(\bar{t}_s)_{opt}} = \frac{M}{4} \sqrt{\frac{1}{D_{1D} \tau_{1D}}} = \frac{M}{\bar{n}}$$

  Facilitated diffusion speeds up the search over pure 1D diffusion by a factor or M/$\bar{n}$.

$$M \approx 5\text{x}10^6 \text{ bp}$$
$$\bar{n} \approx 200 - 500 \text{ bp}$$
$$\text{Optimal facilitated diffusion is} \sim 10^2 \text{faster than 3D}$$
$$\sim 10^4 \text{faster than 1D}$$

Energetics of Diffusion

What determines the diffusion coefficient for sliding and $\bar{\tau}_1$? We need the non-specific protein interaction to be strong enough that it doesn't dissociate too rapidly, but also weak enough that it can slide rapidly. To analyze this, we use a model in which the protein is diffusing on a modulated energy landscape looking for a low energy binding site.



Model[2]

- Assume each sequence can have different interaction with the protein.
- Base pairs in binding patch contribute additively and independently to give a binding energy E_n for each site, n.
- Assume that the variation in the binding energies as a function of site follow Gaussian random statistics, characterized by the average binding energy $\langle E \rangle$ and the surface energy roughness σ.
- The protein will attempt to move to an adjacent site at a frequency $\nu = \Delta\tau^{-1}$. The rate of jumping is the probability that the attempt is successful times ν, and depends on the energy difference between adjacent sites, $\Delta E = E_{n\pm1} - E_n$. The rate is ν if $\Delta E < 0$, and $\nu \cdot \exp[-\Delta E/kBT]$ for $\Delta E > 0$.

Calculating the mean first passage time to reach a target site at a distance of L base pairs from the original position yields

$$\bar{\tau}_{1D} = L^2 \Delta\tau \left(1 + \frac{1}{2}\left(\frac{\sigma}{k_B T}\right)^2\right)^{-1/2} e^{-7\sigma^2/4(k_B T)^2}$$

Which follows a diffusive form with a diffusion constant

$$D_{1D} = \frac{L^2}{2\bar{\tau}_{1D}} = \frac{1}{2\Delta\tau}\left(1 + \frac{1}{2}\left(\frac{\sigma}{k_B T}\right)^2\right)^{1/2} e^{-7\sigma^2/4(k_B T)^2} \tag{15.3.2}$$

Using this to find conditions for the fastest search time:

$$t_{opt} = \frac{M}{2}\sqrt{\frac{\pi\bar{\tau}_{3D}}{4D_{1D}}} \qquad \bar{n}_{opt} = \sqrt{\frac{16}{\pi} D_{1D}\bar{\tau}_{3D}} \qquad \bar{\tau}_{1D} = \bar{\tau}_{3D}$$

Speed vs Stability Paradox

Speed: Fast speed $\rightarrow$ fast search in 1D. From eq. (15.3.2), we see that

$$D_{1D} \propto \exp\left[-\left(\frac{\sigma^2}{k_B T}\right)\right] \tag{15.3.3}$$

With this strong dependence on σ, effective sliding with proper $\overline{n}$ requires

$$\sigma < 2k_B T$$

Stability: On the other hand, we need to remain stably bound for proper recognition and activity. To estimate we argue that we want the equilibrium probability of having the protein bound at the target site be $P_{eq} \approx 0.25$. If E_0 is minimum energy of the binding site, and the probability of occupying the binding site is the following. First we can estimate that

$$E_0 \approx -\sigma\sqrt{2\log M}$$

which suggests that for adequate binding:

$$\sigma > 5k_B T$$

Proposed Two-State Sliding Mechanism

To account for these differences, a model has been proposed:

- While 1D sliding, protein is constantly switching between two states, the search and recognize conformations: $S \rightleftharpoons R$. S binds loosely and allows fast diffusion, whereas R interacts more strongly such that σ increases in the R state.
- These fast conformational transitions must have a rate faster than

$$> \frac{\overline{n}}{\overline{\tau}_{1D}} \sim 10^4 s^{-1}$$

- Other Criteria:

$$\langle E_R \rangle < \langle E_S \rangle$$
$$\sigma_R > \sigma_S$$

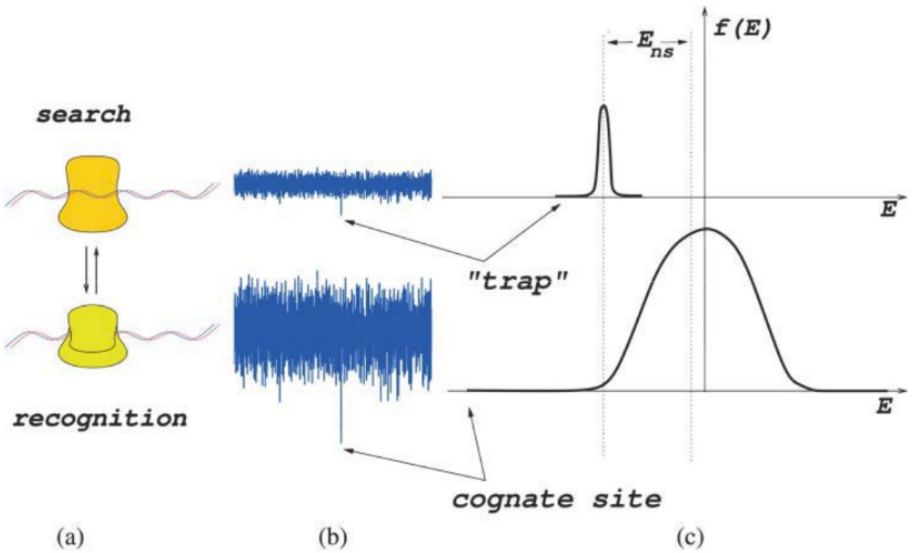


FIGURE 6 Cartoon demonstrating the two-mode search-and-fold mechanism. (*Top*) Search mode; (*bottom*) recognition mode. (*a*) Two conformations of the protein bound to DNA: partially unfolded (*top*) and fully folded (*bottom*). (*b*) The binding energy landscape experienced by the protein in the corresponding conformations. (*c*) The spectrum of the binding energy determining stability of the protein in the corresponding conformations.

	Reprinted from M. Slutsky and L. A. Mirny, Kinetics of protein-DNA interaction: Facilitated target location in sequence-dependent potential, Biophys. J. 87 (6), 4021–4035 (2004), with permission from Elsevier.

Diffusion on rough energy landscape

The observation in eq. (15.3.3), relating the roughness of an energy landscape to an effective diffusion rate is quite general.[3] If we are diffusing over a distance long enough that the corrugation of the energy landscape looks like Gaussian random noise with a standard deviation σ, we expect the effective diffusion coefficient to scale as

$$D_{eff} = D_0 \exp\left[-\left(\frac{\sigma^2}{k_B T}\right)\right] \tag{15.3.4}$$

where D_0 is the diffusion constant in the absence of the energy roughness.

Single-Molecule Experiments

To now there still is no definitive evidence for coupled 1D + 3D transport, although there is a lot of data now showing 1D sliding. These studies used flow to stretch DNA and followed the position of fluorescently labelled proteins as they diffused along the DNA.

Austin: Lac Repression follow up → observed D_{1D} varies by many orders of magnitude.[4]

$$D_{1D} : 10^2 - 10^5 \text{ nm}^2/\text{s}$$
$$\overline{n} \approx 500 \text{ nm}$$

Blainey and Xie: hOGG1 DNA repair protein:[5]

$$\Delta G^{\dagger}_{slide} \approx 0.5 \text{kcal/mol} \approx k_B T$$
$$D_{1D} \sim 5\text{x}10^6 \text{ bp}^2/\text{s}$$
$$\overline{n} \approx 440 \text{ bp}$$

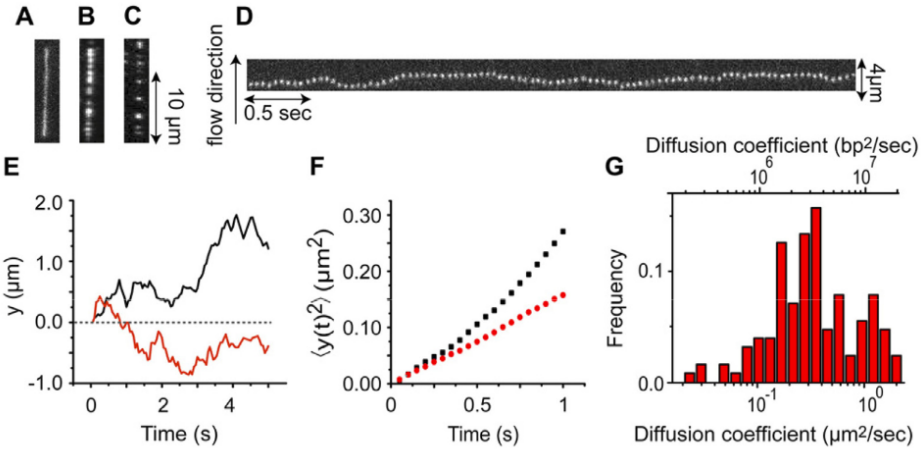


Figure 7. A) Kymograph of an individual fluorescently labeled p53 transcription factor moving along flow-stretched DNA. The *x* axis represents time and the flow is directed upward along the *y*-axis. B) Trajectories of two p53 proteins diffusing on λ-DNA. C) Mean square displacement (MSD) versus time of the same two trajectories. D) Histogram of the diffusion coefficient *D* of 162 individual p53 proteins. Figure reproduced with permission from

	Reprinted from A. Tafvizi, F. Huang, J. S. Leith, A. R. Fersht, L. A. Mirny and A. M. van Oijen, Tumor Suppressor p53 Slideson DNA with Low Friction and High Stability, Biophys. J. 95 (1), L01–L03 (2008), with permission from Elsevier.

$$D_{1D} \qquad 10^6 - 10^7 \text{ bp}^2/\text{s} \approx 10^{-1} - 10^0 \ \mu\text{m}^2/\text{s}$$

1. (5)M. Slutsky and L. A. Mirny, Kinetics of protein-DNA interaction: Facilitated target location in sequence-dependent potential, Biophys. J. 87 (6), 4021–4035 (2004); A. Tafvizi, L. A. Mirny and A. M. van Oijen, Dancing on DNA: Kinetic aspects of search processes on DNA, Chemphyschem 12 (8), 1481–1489 (2011).
2. M. Slutsky and L. A. Mirny, Kinetics of protein-DNA interaction: Facilitated target location in sequence-dependent potential, Biophys. J. 87 (6), 4021–4035 (2004).
3. R. Zwanzig, Diffusion in a rough potential, Proc. Natl. Acad. Sci. U. S. A. 85 (7), 2029 (1988).
4. Y. M. Wang, R. H. Austin and E. C. Cox, Single molecule measurements of repressor protein 1D diffusion on DNA, Phys. Rev. Lett. 97 (4), 048302 (2006).
5. P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine and X. S. Xie, A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA, Proc. Natl. Acad. Sci. U. S. A. 103 (15), 5752 (2006).