15. Passive Transport

Passive transport is often synonymous with diffusion, where thermal energy is the only source of motion.

$$\langle r(t) \rangle = 0$$
 $\langle r^2(t) \rangle^{1/2} = \sqrt{6Dt}$ $r_{rms} \propto \sqrt{t}$

In biological systems, diffusive transport may work on a short scale, but it is not effective for long-range transport. Consider:

$$\langle r^2 \rangle^{1/2}$$
 for small protein moving in water
~10 nm $\rightarrow 10^{-7}$ s
~10 μ m $\rightarrow 10^{-1}$ s

Active transport refers to directed motion:



This requires an input of energy into the system, however, we must still deal with random thermal fluctuations.

How do you speed up transport?

We will discuss these possibilities:

- Reduce dimensionality: Facilitated diffusion
- Free energy (chemical potential) gradient: Diffusion in a potential
- Directional: Requires input of energy, which drives the switching between two conformational states of the moving particle tied to translation.





Dimensionality Reduction

One approach that does not require energy input works by recognizing that displacement is faster in systems with reduced dimensionality. Let's think about the time it takes to diffusively encounter a small fixed target in a large volume, and how this depends on the dimensionality of the search. We will look at the mean first passage time to find a small target with radius *b* centered in a spherical volume with radius *R*, where $R \gg b$. If the molecules are initially uniformly distributed within the volume the average time it takes for them to encounter the target (i.e., MFPT) is¹

$$\langle \tau_{3D} \rangle \approx \frac{R^2}{3D_3} \left(\frac{R}{b} \right) \qquad R \gg b$$

$$\langle \tau_{2D} \rangle \approx \frac{R^2}{2D_2} \ln \left(\frac{R}{b} \right) \qquad R \gg b$$

$$\langle \tau_{1D} \rangle \approx \frac{R^2}{3D_1}$$

Here D_n is the diffusion constants in *n* dimensions (cm²/sec). If we assume that the magnitude of *D* does not vary much with *n*, the leading terms in these expressions are about equal, and the big differences are in the last factor

$$\left(\frac{R}{b}\right) > \ln\left(\frac{R}{b}\right) \gg 1$$
$$\left\langle \tau_{3D} \right\rangle > \left\langle \tau_{2D} \right\rangle \gg \left\langle \tau_{1D} \right\rangle$$

Based on the volume that needs searching, there can be a tremendous advantage to lowering the dimensionality.

O. G. Berg and P. H. von Hippel, Diffusion-controlled macromolecular interactions, Annu. Rev. Biophys. Biophys. Chem. 14, 131-158 (1985); H. C. Berg and E. M. Purcell, Physics of chemoreception, Biophys. J. 20, 193-219 (1977).

Facilitated Diffusion²

Facilitated diffusion is a type of dimensionality reduction that has been used to describe the motion of transcription factors and regulatory proteins looking for their binding target on DNA.

E. coli Lac Repressor

Experiments by Riggs et al. showed that *E. coli* Lac repressor finds its binding site about one hundred times faster than expected by 3D diffusion.³ They measured $k_a=7\times10^9 \text{ M}^{-1} \text{ s}^{-1}$, which is 100–1000 times faster than typical rates. The calculated diffusion-limited association rate from the Smoluchowski equation is $k_a \approx 10^8 \text{ M}^{-1} \text{ s}^{-1}$ using estimated values of $D\approx5\times10^{-7} \text{ cm}^2 \text{ s}^{-1}$ and $R\approx5\times10^{-8}$ cm. Berg and von Hippel theoretically described the possible ways in which nonspecific binding to DNA enabled more efficient one-dimensional motion coupled to three-dimensional transport.⁴

Many Possibilities for Locating Targets Diffusively: Coupled 1D + 3D Diffusion

- 1) Sliding (1D diffusion along chain as a result of nonspecific interaction)
- 2) Microhop (local translocation with free diffusion)
- 3) Macrohop (...to distal segment via free diffusion)
- 4) Intersegmental transfer at crossing—varies with DNA dynamics



^{2.} P. H. von Hippel and O. G. Berg, Facilitated target location in biological systems, J. Biol. Chem. **264** (2), 675–678 (1989).

A. D. Riggs, S. Bourgeois and M. Cohn, The lac represser-operator interaction, J. Mol. Biol. 53 (3), 401–417 (1970);
 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).

^{4.} O. G. Berg, R. B. Winter and P. H. Von Hippel, Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory, Biochemistry **20** (24), 6929–6948 (1981).

Consider Coupled Sliding and Diffusion: The Steady-State Solution

The transcription factor diffuses in 1D along DNA with the objective of locating a specific binding site. The association of the protein and DNA at all points is governed by a nonspecific interaction. Sliding requires a balance of nonspecific attractive forces that are not too strong (or the protein will not move) or too weak (or it will not stay bound). The nonspecific interaction is governed by an equilibrium constant and exchange rates between the bound and free forms:

$$F \underbrace{\frac{k_a}{k_d}}_{K_d} B \qquad \qquad K = \frac{k_a}{k_d} = \frac{\overline{\tau}_{1D}}{\overline{\tau}_{3D}}$$

We can also think of this equilibrium constant in terms of the average times spent diffusing in 1D or 3D. The protein stays bound for a period of time dictated by the dissociation rate k_d . It can then diffuse in 3D until reaching a contact with DNA again, at a point which may be short range in distance but widely separated in sequence.



The target for the transcription factor search can be much larger that the physical size of the binding sequence. Since the 1D sliding is the efficient route to finding the binding site, the target size is effectively covered by the mean 1D diffusion length of the protein, that is, the average distance over which the protein will diffuse in 1D before it dissociates. Since one can express the average time that a protein remains bound as $\overline{\tau}_{1D} = k_d^{-1}$, the target will have DNA contour length of

$$R^* = \left(\frac{4D_1}{k_d}\right)^{1/2}$$

If the DNA is treated as an infinitely long cylinder with radius b, and the protein is considered to have a uniform probability of nonspecifically associating with the entire surface of the DNA, then one can solve for the steady-state solution for the diffusion equation, assuming a completely absorbing target. The rate constant for specific binding to the target has been determined as

$$\eta = \frac{D_1 K'}{D_3 b}$$

where K' is the equilibrium constant for nonspecific binding per unit surface area of the cylinder $(M^{-1} \text{ cm}^{-2} \text{ or cm})$. We can express the equilibrium constant per base-pair as $K = 2\pi \ell b K'$, where ℓ is the length of a base pair along the contour of the DNA. The association rate will be given by the product of k_{bind} and the concentration of protein.

Search Times in Facilitated Diffusion⁵

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

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

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 \overline{n} , the average number of cycles $\overline{k} = M/\overline{n}$, and the average search time can be written as

$$\overline{t}_{s} = \frac{M}{\overline{n}} \left(\overline{\tau}_{1D} + \overline{\tau}_{3D} \right) \tag{1}$$

 $\overline{\tau}$ is the mean search time during one cycle. If we assume that sliding occurs through normal 1D diffusion, then we expect that $\overline{n} \propto \sqrt{D_{1D}\overline{\tau}_{1D}}$, where the diffusion constant is expressed in units of bp²/s. More accurately, it is found that if you executed a random walk with an exponentially weighted distribution of search times:

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

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

$$\frac{\partial \overline{t_s}}{\partial \tau_{1D}} = 0$$

and find that t_{ont} corresponds to the condition

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

Using this in eq. (1) we have

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

M. Slutsky and L. A. Mirny, Kinetics of protein-DNA interaction: Facilitated target location in sequencedependent 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).

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

• **3D only:** $\overline{\tau}_{1D} \to 0 \quad \therefore \overline{n} \to 1$ leading to

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

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.

$$\frac{\overline{t_{3D}}}{\left(\overline{t_s}\right)_{opt}} = \frac{\overline{n}}{2}$$

• **1D only:** $\overline{\tau}_{3D} \to 0 \quad \therefore \overline{n} \to M$, and

$$\overline{t}_{\rm ID} \approx \frac{M^2}{4D_{\rm ID}}$$
$$\frac{\overline{t}_{\rm ID}}{\left(\overline{t}_s\right)_{\rm opt}} = \frac{M}{4} \sqrt{\frac{1}{D_{\rm ID}\tau_{\rm ID}}} = \frac{M}{\overline{n}}$$

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

Example: Bacterial Genome

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

Energetics of Diffusion

What determines the diffusion coefficient for sliding and $\overline{\tau}_1$? 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.



Model⁶

- 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 (E) and the surface energy roughness σ.
- The protein will attempt to move to an adjacent site at a frequency v = Δτ⁻¹. The rate of jumping is the probability that the attempt is successful times v, and depends on the energy difference between adjacent sites, ΔE=E_{n±1}-E_n. The rate is v if ΔE<0, and v·exp[-ΔE/k_BT] for ΔE>0.

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

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

Which follows a diffusive form with a diffusion constant

M. Slutsky and L. A. Mirny, Kinetics of protein-DNA interaction: Facilitated target location in sequencedependent potential, Biophys. J. 87 (6), 4021–4035 (2004).

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

Using this to find conditions for the fastest search time:

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

Speed vs Stability Paradox

Speed: Fast speed \rightarrow fast search in 1D. From eq. (2), we see that

$$D_{\rm 1D} \propto \exp\left[-\left(\frac{\sigma^2}{k_B T}\right)\right]$$
 (3)

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

 $\sigma < 2k_{\rm B}T$

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:

 $\sigma > 5k_BT$

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 sequencedependent potential, Biophys. J. 87 (6), 4021–4035 (2004), with permission from Elsevier.

Diffusion on rough energy landscape

The observation in eq. (3), relating the roughness of an energy landscape to an effective diffusion rate is quite general.⁷ 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

$$D_{eff} = D_0 \exp\left[-\left(\frac{\sigma^2}{k_B T}\right)\right]$$
(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 \rightarrow observed D_{1D} varies by many orders of magnitude.⁸

$$D_{\rm 1D}: 10^2 - 10^5 \text{ nm}^2/\text{s}$$

 $\overline{n} \approx 500 \text{ nm}$

Blainey and Xie: hOGG1 DNA repair protein:9

^{7.} R. Zwanzig, Diffusion in a rough potential, Proc. Natl. Acad. Sci. U. S. A. 85 (7), 2029 (1988).

^{8.} 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).





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 Slides on DNA with Low Friction and High Stability, Biophys. J. **95** (1), L01–L03 (2008), with permission from Elsevier.

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

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).