MAE 545: Lecture 19 (4/27) Figure 1. (*A*) Schematic representation of the protein–DNA search problem. The protein (yellow)

How proteins find target sites on DNA? or cell nucleus (in eukaryotes). Compare with figure 9(*A*) which shows confined DNA. (*B*) The target site must be recognized with 1 and 2 nm precision, as displacement by 1 and 2 pp results by 1 bp results by 1 \sim

Statistical mechanics of polymers

 10 nm

Figure 2. (*A*) The mechanism of facilitated diffusion. The search process consists of alternating **rounds of 3D and 1D and 1D and 1D and 1D diffusion, each with a non-age duration to 1D, respectively. (Growth dynamics of actin** This allows the protein to associate some distance ∼*n*¯ away from the target site and reach it by sliding, effectively increasing the reaction cross-section from 1bp to ∼*n*¯. The antenna effect is **filaments and microtubules**

Berg - von Hippel theory (1980s) or cell nucleus (in eukaryotes). Compare with figure 9(*A*) which shows confined DNA. (*B*) The target site must be recognized with 1 base-pair (0*.*34 nm) precision, as displacement by 1 bp results

(facilitated diffusion)

- **1. Proteins diffuse in space and nonspecifically bind to a random location on DNA.**
- **2. Proteins slide (diffuse) along the DNA.**
- **3. Proteins jump (diffuse) to another random location on DNA and continue this sliding/ jumping process until the target site is found.**

Figure 2. (*A*) The mechanism of facilitated diffusion. The search process consists of alternating D_3 - diffusion constant in space antenna effect [9]. During 1D diffusion (sliding) along DNA, a protein visits on average *n*¯ sites. D_1 − diffusion constant along the DNA $b = 0.34$ nm L **- DNA length**

How long that is it take to find a target site in this process? *1.3. History of the problem: theory*

2

O.G.Berg et al., Biochemistry **20**, 6929-48 (1981) 3D diffusion and effectively 1D diffusion of protein along DNA (*the 1D/3D mechanism*) was suggested. This mechanism was first proposed and dismissed by Riggs *et al* [1] but was soon $PIOCIVIVIUU$ \blacktriangleright \blacktriangleright </u>

Berg - von Hippel theory (1980s) target site must be recognized with 1 base-pair (0*.*34 nm) precision, as displacement by 1 bp results

First assume fixed sliding time τ_{1d}

Number of distinct sites visited during each sliding event

 $n = \sqrt{16D_1\tau_{1d}/(\pi b^2)}$

(valid for *n*>>1)

Probability that target site is found during a sliding event

$$
q=nb/L
$$

Probability that target site is found exactly after N_R rounds

 $p(N_R) = q(1-q)^{N_R-1}$

Average number of rounds needed to find the target ∞

$$
\overline{N_R} = \sum_{N_R=1} N_R p(N_R) = 1/q
$$

Figure 2. (*A*) The mechanism of facilitated diffusion. The search process consists of alternating $b = 0.34$ nm L **- DNA length**

rounds of 3D and 1D diffusion, each with average duration τ3D and τ1D, respectively. (*B*) The antenna effect [9]. During 1D diffusion (sliding) along DNA, a protein visits on average *n*¯ sites. D_3 - diffusion constant in space

- diffusion constant along the DNA sliding, effectively increasing the reaction cross-section from 1bp to ∼*n*¯. The antenna effect is D_1 - diffusion constant along the DNA

responsible for the speed-up tend-up the speed-up tend-up to the speed-up tend-up to the speed-up tend-up tend
in a characteristic iumning time τ_{3d} - characteristic jumping time

Average search time *1.3. History of the problem: theory*

$$
\boxed{\overline{t}_s = \overline{N_R} \left(\tau_{1d} + \tau_{3d} \right)}
$$

O.G.Berg et al., Biochemistry **20**, 6929-48 (1981) \cap C Rora ot al by Berg and Blomberg [4] and finally developed by Berg *et al* [5]. The basic idea of the 1D/3D $proontning$ **</u>**

suggested. This mechanism was first proposed and dismissed by Riggs *et al* [1] but was soon

Facilitated diffusion in a different sequence and consequently a different site.

In reality sliding times are exponentially distributed

$$
p(\tau_{1d}) = k_{\text{off}}^{\text{NS}} e^{-k_{\text{off}}^{\text{NS}} \tau_{1d}}
$$

$$
\langle \tau_{1d} \rangle = \int_0^\infty d\tau_{1d} \,\tau_{1d} \, p(\tau_{1d}) = 1/k_{\text{off}}^{\text{NS}}
$$

Average number of distinct sites visited during each sliding

$$
\langle n \rangle = \int_0^\infty d\tau_{1d} p(\tau_{1d}) \sqrt{16D_1 \tau_{1d}/(\pi b^2)}
$$

$$
\langle n \rangle = 2\sqrt{D_1 \langle \tau_{1d} \rangle / (b^2)}
$$

Average probability that target site is found during a sliding event

 $\langle q \rangle = \langle n \rangle b/L$

Average number of rounds *NR* **needed to find the target site**

$$
\overline{\langle N_R\rangle}=1/\left\langle q\right\rangle
$$

Figure 2. (*A*) The mechanism of facilitated diffusion. The search process consists of alternating rounds of 3D and 1D and 1D and 1D and 1D and 1D, respectively. The state of the time termine the termine termin
F diffusion constant in space antenna effect [9]. During 1D diffusion (sliding) along DNA, a protein visits on average *n*¯ sites. **- diffusion constant along the DNA** sliding, effectively increasing the reaction cross-section from 1bp to ∼*n*¯. The antenna effect is responsible for the speed-up tend-up the speed-up tend-up to the speed-up tend-up to the speed-up tend-up tend
In a characteristic iumning time $b = 0.34$ nm L **- DNA length** D_3 - diffusion constant in space D_1 - diffusion constant along the DNA τ_{3d} - characteristic jumping time

Average search time

$$
\overline{\langle t_s \rangle} = \overline{\langle N_R \rangle} (\langle \tau_{1d} \rangle + \tau_{3d})
$$
\n
$$
\overline{\langle t_s \rangle} = \frac{L}{2\sqrt{D_1 \langle \tau_{1d} \rangle}} (\langle \tau_{1d} \rangle + \tau_{3d})
$$

Facilitated diffusion

(1D) Brownian motion or a random walk. Upon dissociation from the DNA, the protein

Example: search time for target site in bacteria on DNA with 106 base pairs

Simultaneous search for target site by multiple proteins must find its target site \mathcal{A} and \mathcal{A} are confined with the cell nucleoid (in backeria) (in backeria) Simultaneous skarch for target site by **Shows** target site must be recognized with 1 base-pair (0,1 bp results) present by 1 bp results, and $\overline{\mathbf{A}}$. At \mathbf{A} and \mathbf{A} at \mathbf{A} and \mathbf{A} **Sinultaneous search for tark** <u>emperant sur de dipiri nuclear</u>

...
Anteractions and collisions \rightarrow interactions and considers \rightarrow or cell nucleus (*in eucare in nored*). Compare with figure 9(*A*) which shows confined DNA. (*B)* $\frac{1}{2}$ the must be recognized with 1 base-pair (0,³⁴ nm) precision, as displacement by 1 bp results and 1 bp results by 1 bp results and 1 bp **Figure 1. (***A***)** Search proteine are ignored. The problem of the problem. The pro must find its target site (red) on a long DNA molecule confined within the cell nucleoid (in bacteria) must find its target site (red) on a long DNA molecule confined with \mathcal{L} \Box **letoractions and collisions** Figure 1. (*A*) Schematic representation of the protein (yellow) of the protei the *displacement* with 1 base-pair (1 bp results) **between proteins are ignored Interactions and collisions**

Example 18 Set of the Example 2008 Search times for target site by individual

> $p_1(t_s) = \frac{1}{t_s}$ $\langle t_s \rangle$ $e^{-t_s/\langle t_s \rangle}$

What is the typical search time for the fastest *n n* **of** *n* **independently searching proteins?**

rounds of 3D and 1D and 1D and 1D and 1D and 1D and 1D, respectively. The second terminal terminal terminal te
Property of the terminal termi **(Extreme value distributions)**

by Berg and Blomberg [4] and finally developed by Berg *et al* [5]. The basic idea of the 1D/3D

$$
p_n(t_s) = n \times p_1(t_s) \times \left(\int_{t_s}^{\infty} dt' \ p_1(t') \right)^{n-1} = \frac{n}{\langle t_s \rangle} e^{-nt_s/\overline{\langle t_s \rangle}}
$$

Figure 1. (*A*) Schematic representation of the protein–DNA search problem. The protein (yellow)

1.3. History of the problem: theory probability that one of probabili responsible for the speed-up by facilitated diffusion. responsible for the speed-up by facilitated diffusion. **Figure 2. (***A***)** The meaning of the mechanism of probability that of rounds of 3D and 1D diffusion, each with average duration τ3D and τ1D, respectively. (*B*) The antenna proteins india in diffusion on a proteins take long **namelon** target site at time *t*_s to find the target sliding, effectively increasing the reaction cross-section from 1bp to ∼*n*¯. The antenna effect is This allows the contrains to a
This archives the distribution of a away from the target site and reach it by a more than the target site and r sliding, effectively increasing the reaction cross-section from 1bp to ∼*n*¯. The antenna effect is *n* **proteins finds the**

Figure 2. (*A*) The mechanism of facilitated diffusion. The search process consists of alternating

sliding, effectively increasing the reaction cross-section from 1bp to ∼*n*¯. The antenna effect is

 $\mathsf n$ proteins finds the uproteins take longer time 3D diffusion and effectively 1D diffusion of protein along DNA (*the 1D/3D mechanism*) was suggsted. This was first proposed and dismissed by Riggs **proposed and dispite by Riggs and All but was soon Figure 2.** (*A*) The mechanism of facilitated diffusion. The search process consists of alternating rounds of 3D and 1D diffusion, each with average duration τ3D and τ1D, respectively. (*B*) The n proteins finds the proteins take longer time $\overline{}$ **target site at time t's a to find the target site and reach it by** sliding, effectively increasing the reaction cross-section from 1bp to ∼*n*¯. The antenna effect is sliding, effectively increasing the reaction cross-section from 1bp to ∼*n*¯. The antenna effect is antenna effect [9]. During 1D diffusion (sliding) along DNA, a protein visits on average *n*¯ sites. This appropendium to a target some distance and reach and reach and reach it and reach it and reach it and reach i antenna effect [9]. During 1D diffusion (sliding) along DNA, a protein visits on average *n*¯ sites. This about the protein to a about the control and reach it b probability that one of probability that other n-1

This allows the protein to associate some distance ∞nd reach it and reach it by a protein the target site and reach it by a protein the target site and reach it by a protein the target site and reach it by a protein the t **rounds of 3D and 1D diffusion, each with average duration, each with average duration to 3D and the 1D, respectively. (rounds and 1D and 1D and 1D average search time is and time is and the search of the is and the search time is and the search tim reduced by factor** *n*

n proteins finds the **proteins take longer time**
target site at time *t*_s to find the target site

$$
\int_0^\infty dt_s \, t_s \, p_n(t_s) = \frac{\overline{\langle t_s \rangle}}{n}
$$

Statistical mechanics of polymers and filaments

Statistical mechanics of polymers and filaments

molecular dynamics simulation

Note: averaging over time is equivalent to averaging over all possible configurations weighted with Boltzmann weights!

partition function (sum over all possible configurations)

expected value of $\langle \emptyset \rangle = \sum_{c}$

$$
Z = \sum_{c} e^{-E_c/k_B T}
$$

$$
\langle \mathcal{O} \rangle = \sum_{c} \mathcal{O}_{c} \frac{e^{-E_{c}/k_{B}T}}{Z}
$$

 E_c **energy of a given configuration**

- *T* **temperature**
- k_B **Boltzmann constant**

 $k_B = 1.38 \times 10^{-23}$ JK⁻¹

Persistence length

correlations between tangents

$$
\langle \mathbf{t}(s) \cdot \mathbf{t}(s+x) \rangle = e^{-x/\ell_p}
$$

tangents become uncorrelated beyond persistence length!

persistence length

$$
\ell_p = \frac{B}{k_B T}
$$

- *B* **filament bending rigidity**
- *T* **temperature**
- *kBT ^L* **filament length**

Short filaments remain straight

 $L \ll l_p$ $L \gg l_p$

Long filaments perform self-avoiding random walk

Examples: persistence length

B

 $k_B T$

 ℓ_p

11

 10_{nm}

Persistence length for polymers is on the order of nm

actin $\ell_p \approx 17 \,\mu \mathrm{m}$

 m icrotubule $\,\ell_p \approx 1.4 \, \mathrm{mm}$

double stranded DNA $\ell_p \approx 50$ nm **single stranded DNA**

 $\ell_p \approx 2$ nm

uncooked spagetthi

 $\ell_p \approx 10^{18}$ m

End-to-end distance

Ideal chain vs worm-like chain

N **identical unstretchable links (Kuhn segments) of length** *a* **with freely rotating joints**

Each configuration *C* **has zero energy cost.**

$$
E_c=0
$$

Ideal chain Worm-like chain

Continuous unstretchable rod

Bending energy cost of configuration *C***:**

$$
E_c=\frac{B}{2}\int_0^L ds \left(\frac{d^2\vec{r}}{ds^2}\right)^2
$$

Each configuration *C* **appears with probability** $p_c \propto e^{-E_c/k_B T}$

L **=** *Na* **- chain length**

Ideal chain vs worm-like chain

N **identical unstretchable links (Kuhn segments) of length** *a* **with freely rotating joints**

Ideal chain Worm-like chain

Continuous unstretchable rod

$$
\left<\vec{R}_{AB}^2\right>=Na^2=aL
$$

 $\overline{}$ \vec{R}_{\angle}^2 $\binom{2}{AB}$ $\approx 2\ell_p L =$ 2*BL* $k_{B}T$

End-to-end distance fluctuations can be made identical if one choses the segment length to be

$$
a=2\ell_p
$$

L **=** *Na* **- chain length**

Stretching of ideal freely jointed chain

Exact result for end-to-end distance

 \hat{x}

 \hat{y}

*z*ˆ

Stretching of worm-like chains

Assume long chains $L \gg \ell_p$

$$
\frac{F\ell_p}{k_B T} = \frac{1}{4} \left(1 - \frac{\langle x \rangle}{L} \right)^{-2} - \frac{1}{4} + \frac{\langle x \rangle}{L}
$$

J.F. Marko and E.D. Siggia,

Macromolecules **28**, 8759-8770 (1995)

Experimental results for stretching of DNA 8760 Marko and Siggia *Macromolecules, Vol. 28, No. 26, 1995*

 $L = 32.8 \mu m$

Random coil to globule transition in polymers $T > \Theta$ *T* < Θ **random coil compact globule**

$$
R \sim \sqrt{L\ell_p} \qquad R \sim \big(
$$

at high temperature entropic contributions dominate

$$
R\sim \left(d^2L\right)^{1/3}
$$

at low temperature strong, first-order-like, coil-to-native transitions. attraction between polymer model show the transition from the transition from the coil to the coil to the coil to the coil to the native station by the nuclear barrier. d - diameter of polymer chain interaction parameters single singl of, usually symmetry-related, lowest-energy states). Hence, the nearest-neighbor interaction in the Goj model is similar to *E*p: (9) Doniach, S., Garel, T.; Orland, H. *J. Chem. Phys.* **1996**, *105*, 1605. attraction between polymer chains dominates **Figure 6.** Snapshot showing the molten-globule state of a 512-segment semiflexible lattice polymer at *B*/*E*^p) 0.1 and *k*B*T*/*E*^p) 3.289 (symbol **at low temperature**

A. <u>J. Phys. Chem. B</u> 110, 3734 (2006) B I, <u>J. Phys. Chem. t</u> Figures from: W.B. Hu and D. Frenkel, J. Phys. Chem. B 110, 3734 (2006)

Further reading

Dynamics of actin filaments and microtubules

Cytoskeleton in cells

Cytoskeleton matrix gives the cell shape and mechanical resistance to deformation.

Microtubule

(wikipedia)

Crawling of cells

migration of skin cells during wound healing

spread of cancer cells during metastasis of tumors

amoeba searching for food $v \sim 0.1 \mu m/s$

Immune system: neutrophils chasing bacteria

David Rogers, 1950s

Movement of bacteria

Listeria monocytogenes **moving in infected cells**

Julie Theriot (speeded up 150x)

 $v \sim 0.1 - 0.3 \mu m/s$

L. A. Cameron *et al.*, Nat. Rev. Mol. Cell Biol. **1**, 110 (2000)

Molecular motors

A*.*B*.* Kolomeisky, J. Phys.: Condens. Matter **25**, 463101 (2013)

Transport of large molecules around cells (diffusion too slow) $v \sim 1 \mu m/s$

Contraction of muscles

Harvard BioVisions

Cell division

Segregation of chromosomes Contractile ring divides the cell in two

Microtubules Actin

Swimming of sperm cells

Swimming of Chlamydomonas (green alga)

Jeff Guasto

Jeff Guasto

 $v \sim 50 \mu \text{m/s}$ *v* $\sim 60 \mu \text{m/s}$

Bending is produced by motors walking on neighboring microtubule-like structures

Actin filaments

Actin filament growing against the barrier

work done against the barrier for insertion of new monomer

$$
W = Fa
$$

effective monomer free energy potential without barrier

effective monomer free energy potential with barrier

Actin filament growing against the barrier

Maximal force that can be balanced by growing filament k_{or}^+ **(stall force)**

$$
v^{+}(F_{\max}) = 0 \qquad \longrightarrow \qquad F_{\max} = \frac{k_B T}{a} \ln\left(\frac{k_{\text{on}}^{+}[M]}{k_{\text{off}}^{+}}\right)
$$

$$
k_{\text{on}}^{+} \sim 10 \mu \text{M}^{-1} \text{s}^{-1}
$$

\n
$$
k_{\text{off}}^{+} \sim 1 \text{s}^{-1}
$$

\n
$$
[M] \sim 10 \mu \text{M}
$$

\n
$$
a \approx 2.5 \text{nm}
$$

\n
$$
F_{\text{max}} \sim 8 \text{pN}
$$

Movement of bacteria

Listeria monocytogenes Bacterium **moving in infected cellsActin polymerization is pushing Caps prevent bacteria further polymerization in comet tails Actin is** 08 16 90 2 **randomly Julie Theriot (speeded up 150x) depolymerized in comet tails** $v \sim 0.1 - 0.3 \mu m/s$ L. A. Cameron *et al.*,

Nat. Rev. Mol. Cell Biol. **1**, 110 (2000)