MAE 545: Lecture 21,22 (4/24) How proteins find target sites on DNA?

Statistical mechanics of polymers



Growth dynamics of actin filaments and microtubules





Dynamics of molecular motors



How proteins find target sites on DNA?





Translation of mRNA



Protein-DNA interactions



$$k_{\rm off}^{\rm S} = A_{\rm s} e^{-\Delta G^{\rm S}/k_B T} \ll k_{\rm off}^{\rm NS} = A_{\rm s} e^{-\Delta G^{\rm NS}/k_B T}$$
$$\frac{k_{\rm off}^{\rm S}}{k_{\rm off}^{\rm NS}} \sim 10^{-6}$$

How long proteins remain bound on DNA?



Probability that protein remains bound for time t and then it unbinds between time t and $t+\Delta t$:

$$k_{\text{off}}\Delta t \times (1 - k_{\text{off}}\Delta t)^{t/\Delta t}$$

$$\lim_{t \to 0} b = k_{\text{off}}e^{-k_{\text{off}}t}$$

Average binding time $\langle t \rangle = \int_0^{1} t p(t) dt = \frac{1}{k_{\text{off}}}$

Proteins remain bound to specific target sites for minutes to hours, while they unbind from nonspecific sites after milliseconds to seconds.



$$\frac{d[\text{P-T}]}{dt} = k_{\text{on}}[\text{P}][\text{T}] - k_{\text{off}}[\text{P-T}]$$

[P-T] concentration of proteins bound to target sites
[P] concentration of free proteins
[T] concentration of empty target sites 6

$$\frac{d[\text{P-T}]}{dt} = (k_{\text{on}}[\text{T}])[\text{P}] \equiv \frac{[\text{P}]}{t_s}$$

characteristic search time

$$t_s = (k_{\rm on}[T])^{-1}$$

How quickly proteins find target sites on DNA?

Characteristic search time via 3D diffusion

$$k_{\rm on} = 4\pi D_3 b$$
 $t_s = (k_{\rm on}[{\rm T}])^{-1}$

1917 Smoluchowski theory

Example: characteristic search time for lac repressor protein in E. coli

 $b \approx 0.34 \text{nm}$ $D_3 \approx 30 \mu \text{m}^2/\text{s}$ [T] ~ 1 per cell ~ $10^{-9}M$ $k_{\text{on}} \sim 10^8 M^{-1} s^{-1}$ $t_s \sim 10s$



Molar concentration $1M = 6 \times 10^{26} \text{m}^{-3}$

in vitro experiments (1970) A. $k_{\rm on}^{\rm exp} \sim 10^{10} M^{-1} s^{-1} \qquad t_s \sim 0.1 s$

A.D.Riggs *et al.*, <u>J. Mol. Biol.</u> **53**, 401-417 (1970)

Why is experimentally observed rate 100 times larger?

Berg - von Hippel theory (1980s)

(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.



 $b = 0.34 \mathrm{nm}$ L - DNA length D_3 - diffusion constant in space D_1 - diffusion constant along the DNA

How long that is it take to find a target site in this process?

8

O.G.Berg et al., <u>Biochemistry</u> **20**, 6929-48 (1981)

Berg - von Hippel theory (1980s)

9

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 $\underset{\infty}{\infty}$

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



b = 0.34nm L - DNA length

- D_3 diffusion constant in space
- $D_1\mbox{-}$ diffusion constant along the DNA
- τ_{3d} characteristic jumping time

Average search time

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

O.G.Berg et al., <u>Biochemistry</u> **20**, 6929-48 (1981)

Facilitated diffusion

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

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

Average probability that target site is found during a sliding event

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

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

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



 $b = 0.34 \mathrm{nm}$ L - DNA length D_3 - diffusion constant in space D_1 - diffusion constant along the DNA au_{3d} - characteristic jumping time

Average search time

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

Facilitated diffusion



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



Sinultaneous search for target site by multiple proteins

Interactions and collisions between proteins are ignored



Search times for target site by individual proteins are exponentially distributed

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

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

(Extreme value distributions)

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

n proteins finds the target site at time *t*_s

probability that one of probability that other n-1 proteins take longer time to find the target site

Average search time is reduced by factor n

 $\int dt_s t_s p_n(t_s) = \frac{\langle t_s \rangle}{n}$

Statistical mechanics of polymers and filaments



Statistical mechanics of polymers and filaments

molecular dynamics simulation



Note: in equilibrium 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 observables

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

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

- $E_c \begin{array}{c} \text{energy of a given} \\ \text{configuration} \end{array}$
 - T temperature
- k_B Boltzmann constant

 $k_B = 1.38 \times 10^{-23} \mathrm{JK}^{-1}$

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**
- L filament length

Short filaments remain straight

 $L \ll \ell_p$

Long filaments perform self-avoiding random walk

 $L \gg \ell_p$



Examples: persistence length

 $\overline{k_BT}$

 ℓ_p



Persistence length for polymers is on the order of nm

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

microtubule $\ell_p \approx 1.4 \,\mathrm{mm}$

└───┘ 10 nm

17

double stranded DNA $\ell_p \approx 50 \,\mathrm{nm}$

single stranded DNA



uncooked spagetthi

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

End-to-end distance



Ideal chain vs worm-like chain

Ideal chain

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



Each configuration *C* has zero energy cost.

$$E_c = 0$$

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 $p_c \propto e^{-E_c/k_BT}$

L = Na - chain length

Ideal chain vs worm-like chain

Ideal chain

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



Worm-like chain

Continuous unstretchable rod



$$\left\langle \vec{R}_{AB}^2 \right\rangle = Na^2 = aL$$

 $\left\langle \vec{R}_{AB}^2 \right\rangle \approx 2\ell_p L = \frac{2BL}{k_B T}$

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



Stretching of worm-like chains

Assume long chains $L \gg \ell_p$



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

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

Macromolecules 28, 8759-8770 (1995)

Experimental results for stretching of DNA

 $L = 32.8 \mu \mathrm{m}$



Random coil to globule transition in polymers
random coilcompact globule
 $T > \Theta$ $T > \Theta$ $T < \Theta$



 $R \sim \sqrt{L\ell_p}$

at high temperature entropic contributions dominate



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

d - diameter of polymer chain at low temperature attraction between polymer chains dominates

Figures from: W.B. Hu and D. Frenkel, <u>J. Phys. Chem. B</u> **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.





(wikipedia)

Crawling of cells



migration of skin cells during wound healing

spread of cancer cells during metastasis of tumors

amoeba searching for food

Immune system: neutrophils chasing bacteria



David Rogers, 1950s $v\sim 0.1 \mu {
m m/s}$

Movement of bacteria

Listeria monocytogenes moving in infected cells



Julie Theriot (speeded up 150x) $v\sim 0.1-0.3 \mu {
m m/s}$



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

Molecular motors



A.B. Kolomeisky, <u>J. Phys.: Condens.</u> <u>Matter</u> **25**, 463101 (2013) **Contraction of muscles**

Transport of large molecules around cells (diffusion too slow)

 $v \sim 1 \mu \mathrm{m/s}$



Harvard BioVisions

https://www.youtube.com/watch?v=FzcTgrxMzZk

Cell division

Segregation of chromosomes

Contractile ring divides the cell in two





Actin

Microtubules

Swimming of sperm cells



Swimming of Chlamydomonas (green alga)



https://sites.tufts.edu/guastolab/movies/

Jeff Guasto

Jeff Guasto

 $v \sim 50 \mu \mathrm{m/s}$

 $v \sim 60 \mu {\rm m/s}$

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 (stall force) $k_{on}^+ \sim$

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

$$k_{\rm on}^+ \sim 10 \mu {\rm M}^{-1} {\rm s}^{-1}$$
$$k_{\rm off}^+ \sim 1 {\rm s}^{-1}$$
$$[M] \sim 10 \mu {\rm M}$$
$$a \approx 2.5 {\rm nm}$$
$$F_{\rm max} \sim 8 {\rm pN}$$

Movement of bacteria



Microtubules



Microtubule dynamic instability







Wikipedia

Simple model of microtubule growth



Typical values in a tubilin solution
of concentration $[T] \approx 10 \mu M$:
 $v_g \approx 2 \mu m/min \quad \propto [T]$
 $v_s \approx 20 \mu m/min \quad \sim const$
 $r_{cat} \approx 0.24 min^{-1} \quad \sim const$
 $r_{res} \approx 3 min^{-1} \quad \propto [T]$

Let's ignore all molecular details and assume that microtubules switch at fixed rates between growing and shrinking phases

Master equation:

 $\frac{\partial p_{\text{growth}}}{\partial t} = -r_{\text{cat}} \, p_{\text{growth}} + r_{\text{res}} \, p_{\text{shrinking}}$ $\frac{\partial p_{\text{shrinking}}}{\partial t} = +r_{\text{cat}} \, p_{\text{growth}} - r_{\text{res}} \, p_{\text{shrinking}}$

$$p_{\rm growth} + p_{\rm shrinking} = 1$$

Steady state ($\partial p / \partial t \equiv 0$):

$$p_{\text{growth}}^* = \frac{r_{\text{res}}}{r_{\text{res}} + r_{\text{cat}}}$$
 $p_{\text{shrinking}}^* = \frac{r_{\text{cat}}}{r_{\text{res}} + r_{\text{cat}}}$

Average growth speed of microtubules

$$\overline{v} = p_{\text{growth}}^* v_g - p_{\text{shrinking}}^* v_s$$

 $\overline{v} \approx 0.4 \,\mu\mathrm{m/min}$

40

How cells control the total length of microtubules



V. Varga et al., Cell 138, 1174-1183 (2009)

Density of motors bound to microtubules



Density of motors bound to microtubules

[*M*] concentration

Time evolution for density of bound motors

$$\frac{\partial \rho(x,t)}{\partial t} = k_{\text{bind}}[M] - v_{\text{mot}} \frac{\partial \rho(x,t)}{\partial x}$$

For initially empty microtubule

$$\rho(x,t) = \begin{cases} \frac{k_{\text{bind}}[M]}{v_{\text{mot}}} x, & 0 < x < v_{\text{mot}}t \\ k_{\text{bind}}[M]t, & x > v_{\text{mot}}t \end{cases}$$

Stationary density of bound motors

$$\rho^*(x) = \frac{k_{\text{bind}}[M]}{v_{\text{mot}}} x$$



Length dependent depolymerization rate



Depolymerization rate is proportional to density of Kip3 motors

$$\rho^*(L) = \frac{k_{\text{bind}}[M]}{v_{\text{mot}}}L$$

V. Varga et al., Nat. Cell Biol. 8, 957-962 (2006)

Controlled length of microtubules



Molecular motors



A.B. Kolomeisky, <u>J. Phys.: Condens.</u> <u>Matter</u> **25**, 463101 (2013) **Contraction of muscles**

Transport of large molecules around cells (diffusion too slow)

 $v \sim 1 \mu \mathrm{m/s}$



Harvard BioVisions

https://www.youtube.com/watch?v=FzcTgrxMzZk

Movement of molecular motors is powered by ATP molecules

Myosin motor walking on actin in muscles

Kinesin motor walking on microtubule



Graham Johnson

https://www.youtube.com/watch?v=oHDRIwRZRVI

https://www.youtube.com/watch?v=YAva4g3Pk6k

Molecular motors vs Brownian ratchets

Myosin motor

ATP driven process drives molecular motors along the filaments

Brownian ratchet

net movement of particles is achieved by periodic modulation of asymmetric external potential





ATP concentration dependent speed of motors



Kinesin motor on microtubules



Maximal speed

 $v_{\rm max} \approx 0.6\,\mu{\rm m/s}$

ATP concentration at half the maximal speed $K_d \approx 50 \,\mu {
m M}$

R. Phillips et al., Physical Biology of the Cell

Motors carrying the load

50

Force exerted on kinesin motors carrying plastic beads can be controlled with optical tweezers



 $F \approx k \Delta x$

Effective spring constant *k* depends on the bead size, refractive indices of the bead and surrounding medium, and the gradient of laser intensity

laser beam

How motor speed depends on the loading force?

K. Visscher et al., <u>Nature</u> **400**, 184-189 (1999)

Motor velocity dependence on the load

kinesin walking on microtubules





How important is viscous drag for motors carrying vesicles?

 $F_{\rm drag} = 6\pi\eta Rv$ $F_{\rm drag} \sim 6\pi 10^{-3} \rm kgm^{-1}s^{-1} \cdot 1\mu m \cdot 1\mu m/s$ $F_{\rm drag} \sim 10^{-2} \rm pN$

Note: viscous drag is negligible

K. Visscher et al., <u>Nature</u> **400**, 184-189 (1999)

ATP concentration dependent stall force



motor step length $a \approx 8 \,\mathrm{nm}$

Position clamp

K. Visscher et al., <u>Nature</u> **400**, 184-189 (1999)



52

Skeletal muscle contraction by myosin motors



Skeletal muscle contraction by myosin motors



Estimated force generated by myosin motors

2pN

Muscles contract at twice the speed of myosin motors

 ~ 0.1 -1 $\mu m/s$

Muscles may contract by 5%-45% per second!

$$300 \times \frac{1}{\pi (30 \text{nm})^2} \sim 20 \text{N/cm}^2$$

0

Skeletal muscle contraction is controlled by nerve cells



Electric signal from nerve cells releases Ca²⁺ from sarcoplasmic reticulum



Low Ca²⁺, muscles are relaxed

(a) Tropomyosin and troponin work together to block the myosin binding sites on actin.



High Ca²⁺, muscles are contracted

(b) When a calcium ion binds to troponin, the troponintropomyosin complex moves, exposing myosin binding sites.



Calcium ion Troponin-tropomyosin complex, moved

How muscles get ATP energy?



How muscles get ATP energy?



How muscles get ATP energy?



Aerobic respiration



Note: Citric acid cycle = Krebs cycle

Electron transport chain



NADH products of the Cytric acid cycle are used to pump H⁺ to the space between outer and inner mitochondrial membrane.

Gradient of H⁺ concentration drives the ATP synthase motor that converts ADP to ATP.

Note: ATP synthase can run in reverse and use ATP to pump H⁺ at low concentrations.

ATP synthase



Energetics of ATP hydrolysis

How much energy is released during ATP hydrolysis?



Chemical potentials are typically defined relative to concentration c₀ ~ 1 M.

 $\mu_s(c_s) = \mu_s(c_0) + k_B T \ln(c_s/c_0)$

Crawling of cells



migration of skin cells during wound healing

spread of cancer cells during metastasis of tumors

amoeba searching for food

Immune system: neutrophils chasing bacteria



David Rogers, 1950s $v\sim 0.1 \mu {
m m/s}$

Crawling of cells

fish skin cell $v = 0.2 \mu m/s$



Swimming of sperm cells

Sperm flagellum is constructed from microtubules





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



Jeff Guasto $v \sim 50 \mu {
m m/s}$

Further reading

