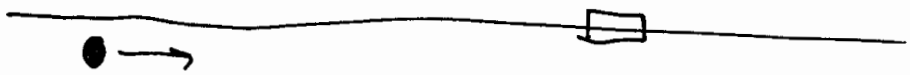
Dynamics of TF - DNA recognition

Search dynamics on <u>smooth</u> landscape

→ assume all sites have G^{NS} except for target

Timescale for thermodynamics to happen?

<u>1D limit</u>



Assume proteins are near to DNA

1D diffusion $N = 5 \times 10^6$ bp ~ 1 mm

$D_{1D} \sim 1\ \mu m^2/sec$ (1/10 of bulk diffusion)

- Search time $T_{1D} \sim \frac{N^2}{D_{1D}} \sim 10^6$ sec

  → too slow

- may be overcome by increasing # TF for P TF molecules

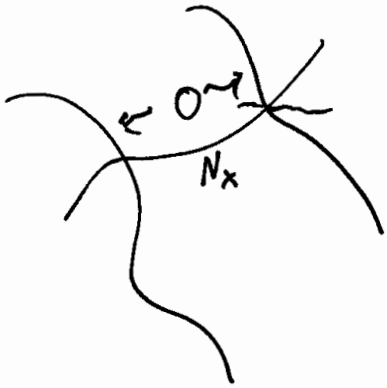
  each TF search over segment of length N

  $T_{1D} \sim \frac{(N/P)^2}{D_{1d}}$ 100 TF's → $T_{1D} \sim 10^2$ sec

  → doable but covers the genome w/ lots of "useless" proteins

  nature tries to keep concentration of unused stuff down

  also things may get too crowded near DNA if there are many different TF's



- time to slide along a segment from one intersection to another

$$T_x \sim \frac{N_x^2}{D_{1d}}$$

- time to search over all segments by 3D random walk

$$T_{comb} \sim \#\text{ segments} \cdot \frac{\text{search time}}{\text{segment}}$$

$$\frac{N}{N_x} \qquad \frac{N_x^2}{D_{1d}}$$

estimate of N_x:

M units of size N_x^3

total length $3 \cdot N_x \cdot M = N = 1\ mm$

all dimension $N_x^3 \cdot M = 1\ \mu m^3$

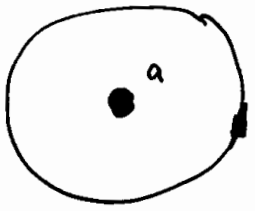
get: $N_x \sim 0.06\ \mu m \ \hat{=}\ 150-200\ bp$

$T_x \sim 3\ msec$

$T_{comb} = 5-10\ sec$

very short search time

Opposite limit : 3D diffusion



$V_{cell} \sim 1\,\mu m^3$ $D_{3d} \sim 10\,\mu m^2/s$

size of binding target $a \sim 1\,nm$

$$T_{3d} \simeq \frac{1}{4\pi} \frac{V_{cell}}{a D_{3d}} \sim 10^6$$

approximately $\frac{V_{cell}}{a^3} \cdot \frac{a^2}{D}$

search volume (# voxels) · search time

for one molecule

→ much faster despite much increased search space

due to reduced redundancy of random walk (random walker does not come back to places it was as often as in 1D)

Realistic model : electrostatic IA confines TF to DNA

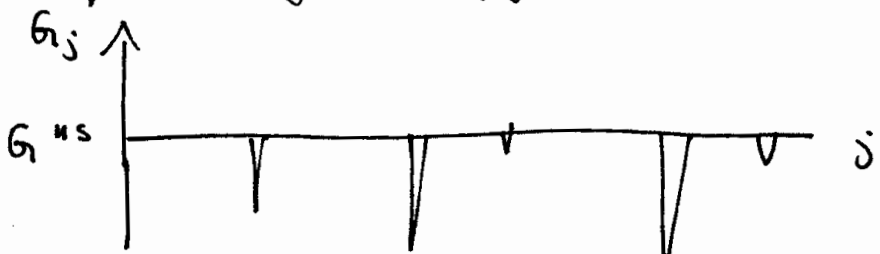
Combined 1D / 3D search

assume dense packing of DNA in cell

- 1D diffusion (sliding) at distance $\leq N_x$
- 3D diffusion at larger distance (hopping between segments)

Include sequence heterogeneity

Consider random walk (locally) on the following energy landscape



assume G^{us} sufficiently small so that only a small fraction of targets have energy $G_j < G^{us}$

- time it take to move from site j to j+1

  if $G_j^{sp} > G^{us}$ then $\tau \equiv \tau_0$ (diffusion)

  if $G_j^{sp} < G^{us}$ then $\tau \equiv \tau_0 \, e^{(G^{us} - G_j^{sp})/kT}$

  thermal activation needed to get out of trap

combine $\tau = \tau_0 \left(1 + e^{(G^{us} - G_j^{sp})/kT}\right)$

- In a typical search, the protein will encounter many such traps, which are uncorrelated with each other in 1D/3D search if # traps / segment is small

Slow down factor due to traps

$$\left\langle e^{(G^{us} - G_i^{sp})/kT} \right\rangle = e^{(G^{us} - G^*)/kT} \cdot \underbrace{\left\langle e^{-\sum_i^L g_i(b_i)/kT} \right\rangle}_{Z^L(\varepsilon)}$$

Slow down factor must not be too large (actually, on order of 1 to match typical expression times)

$$e^{(G^{us} - G^*)/kT} \cdot Z^L(\varepsilon) \sim 1$$

We already require ~~$G^{us} = G^*$~~ $N Z^L(\varepsilon) \sim 1$ from specificity and tunability condition

$$\Rightarrow e^{(G^{us} - G^*)/kT} \sim N$$

$G^{us} - G^*$ large can cause kinetic problem

However, specificity condition also requires

$$N e^{-(G^{us} - G^*)/kT} \leq 1$$

$G^{us} - G^*$ small can cause specificity problem