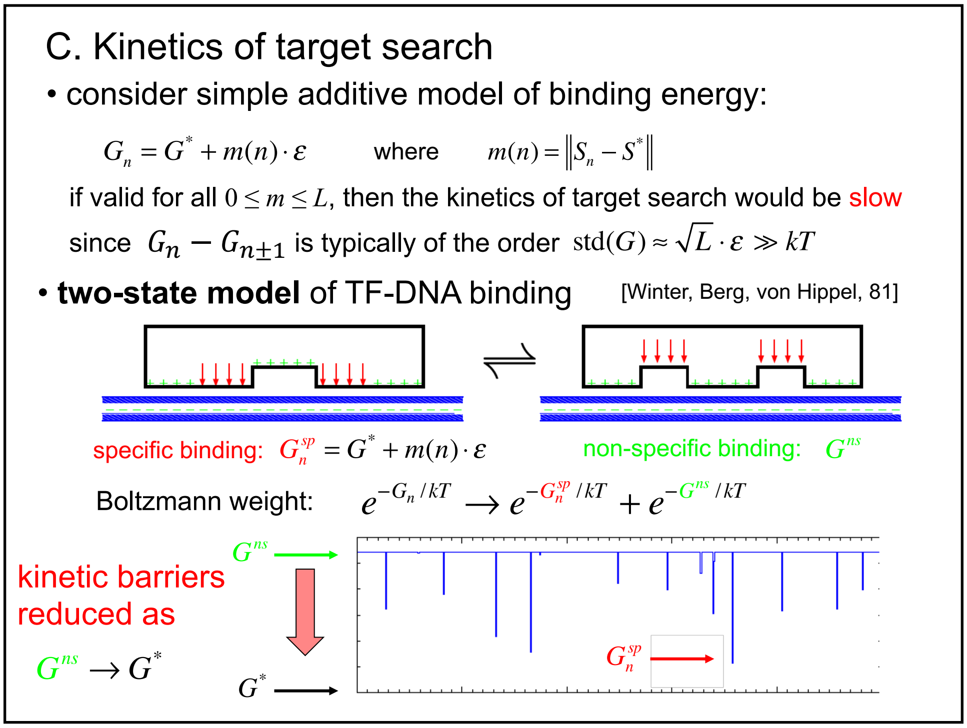


C. Kinetics of target search

- consider simple additive model of binding energy:

$$G_n = G^* + m(n)\cdot\varepsilon \qquad \text{where} \qquad m(n) = \left\|S_n - S^*\right\|$$

if valid for all $0 \le m \le L$, then the kinetics of target search would be slow

since $G_n - G_{n\pm1}$ is typically of the order $\mathrm{std}(G) \approx \sqrt{L}\cdot\varepsilon \gg kT$

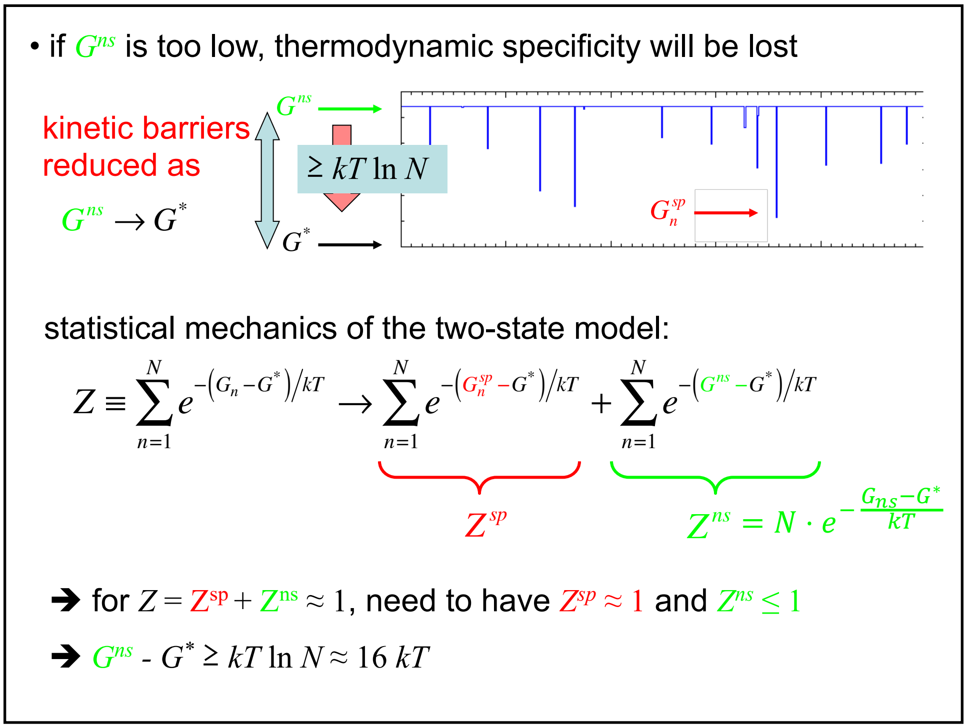
- **two-state model** of TF-DNA binding [Winter, Berg, von Hippel, 81]

specific binding: $G_n^{sp} = G^* + m(n)\cdot\varepsilon$

non-specific binding: G^{ns}

Boltzmann weight: $e^{-G_n/kT} \to e^{-G_n^{sp}/kT} + e^{-G^{ns}/kT}$

kinetic barriers reduced as $G^{ns} \to G^*$



- if G^{ns} is too low, thermodynamic specificity will be lost

kinetic barriers reduced as $G^{ns} \to G^*$

statistical mechanics of the two-state model:

$$Z \equiv \sum_{n=1}^{N} e^{-\left(G_n - G^*\right)/kT} \to \underbrace{\sum_{n=1}^{N} e^{-\left(G_n^{sp} - G^*\right)/kT}}_{Z^{sp}} + \underbrace{\sum_{n=1}^{N} e^{-\left(G^{ns} - G^*\right)/kT}}_{Z^{ns} = N\cdot e^{-\frac{G_{ns}-G^*}{kT}}}$$

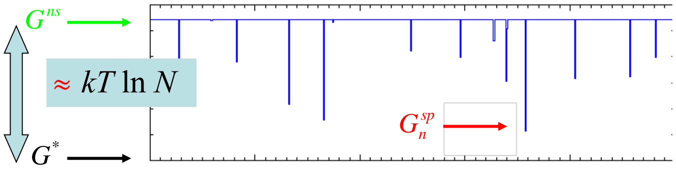
➔ for $Z = Z^{sp} + Z^{ns} \approx 1$, need to have $Z^{sp} \approx 1$ and $Z^{ns} \leq 1$

➔ G^{ns} - $G^* \geq kT \ln N \approx 16\, kT$

- effect of kinetic slow down ?

kinetic barriers reduced as

$G^{ns} \rightarrow G^*$



-- for each trap with binding energy $G^{sp}_n < G^{ns}$

escape time (Kramer): $\tau_n = \tau_0 \cdot e^{\left(G^{ns} - G^{sp}_n\right)/kT}$ τ_0: sliding time

-- average escape time: $$\bar{\tau} \gtrsim \frac{\tau_0}{N} \sum_G \left[1 + e^{(G^{ns} - G)/kT}\right] \cdot \overbrace{\Omega(G)}^{\text{density of state}}$$

$$= \tau_0 \cdot \left[1 + e^{(G^{ns} - G^*)/kT} \cdot Z^{sp}/N\right]$$

➔ for $Z^{sp} \approx 1$, kinetic slowdown insignificant if G^{ns} - $G^* \leq kT \ln N$

➔ both thermodynamics and kinetics okay if G^{ns} - $G^* \approx kT \ln N$

[Note: for the Lac and Arc repressors, G^{ns} - $G^* \approx 15\ kT$]

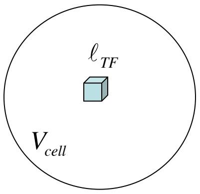
D. Global search dynamics (smooth landscape)

- 1D diffusion along the chromosome: [Elf & Xie, 2007]

$$\left.\begin{array}{l} N = 5 \times 10^6 \text{ bp} \approx 1 \text{ mm} \\ D_1 \approx 0.1\ \mu\text{m}^2 / \text{sec} \end{array}\right\} \quad T_{1D} \sim \frac{N^2}{D_1} \sim 10^7 \text{ sec}$$

-- may be overcome by increasing # of TF; for parallel search, $T_{1D} \sim 1/(\#TF)^2$
-- cost: covers the chromosome with lots of “useless” TFs

- 3D diffusion directly from the cytoplasm:

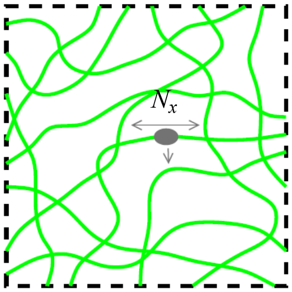


$$T_{3d} \approx \frac{1}{4\pi} \frac{V_{cell}}{\ell_{TF} \cdot D_{cyto}} = \frac{1}{4\pi} \underbrace{(V_{cell}/\ell^3_{TF})}_{\text{search volume}} \cdot \underbrace{(\ell^2_{TF}/D)}_{\text{search time per volume}} \sim 10 \text{ sec}$$

$V_{cell} \approx 3\mu m^3$
$\ell_{TF} \approx 15bp = 5nm$
$D_{cyto} \approx 10\mu m^2 / \text{sec}$

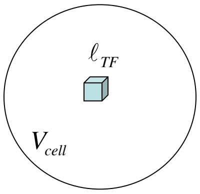
– faster mainly due to the reduced redundancy of 3D random walk
– but TFs typically associate strongly to DNA (sub-compartmentalization)
[e.g., for the Lac repressors, G^{cyto} - $G^{ns} \approx 15\ kT$]

- combined 1D/3D search:
  - dense DNA packing in cell
  - short-time: slide on DNA (over scale N_x)
  - long-time: random walk on 3D network
  - slide dist: $N_\times \sim 300$ bp
  - slide time: $T_\times \sim N_\times^2 / D_1 \sim 0.1$ sec



$$T_{1D/3D} \sim \frac{1}{4\pi}\frac{V_{cell}}{N_\times \cdot (N_\times^2 / T_\times)} \sim 100\,\text{sec}$$

- 3D diffusion directly from the cytoplasm:



$$T_{3d} \approx \frac{1}{4\pi}\frac{V_{cell}}{\ell_{TF}\cdot D_{cyto}} = \frac{1}{4\pi}\underbrace{(V_{cell}/\ell_{TF}^3)}_{\text{search volume}} \cdot \underbrace{(\ell_{TF}^2/D)}_{\text{search time per volume}} \sim 10 \text{ sec}$$

$V_{cell} \approx 3\mu m^3$

$\ell_{TF} \approx 15bp = 5nm$

$D_{cyto} \approx 10\mu m^2 / \text{sec}$

- faster mainly due to the reduced redundancy of 3D random walk
- but TFs typically associate strongly to DNA (sub-compartmentalization)
  [e.g., for the Lac repressors, G^{cyto} - $G^{ns} \approx 15\ kT$]

Summary:

- specificity of target recognition: $Z^{sp} \approx 1$
  - ➔ $\varepsilon \approx 2$ kT, $L \approx 15$ bp, gives $\widetilde{K}_j \equiv \sum_{n\neq j}^{N} e^{(G_j - G_n)/kT} \approx e^{m_j \varepsilon / kT}$
  - ➔ affinity of target sites become "programmable"
- kinetic accessability of target predicts G^{ns} - $G^* \approx 15\ kT$
- combined 1D/3D search

➔ to what extent is "programmable" interactions used ?

➔ search process for multimer?

➔ eukaryotes?

many differences, e.g., $N_p = 10^2 \sim 10^4$ in budding yeast
(need another von Hippel!)