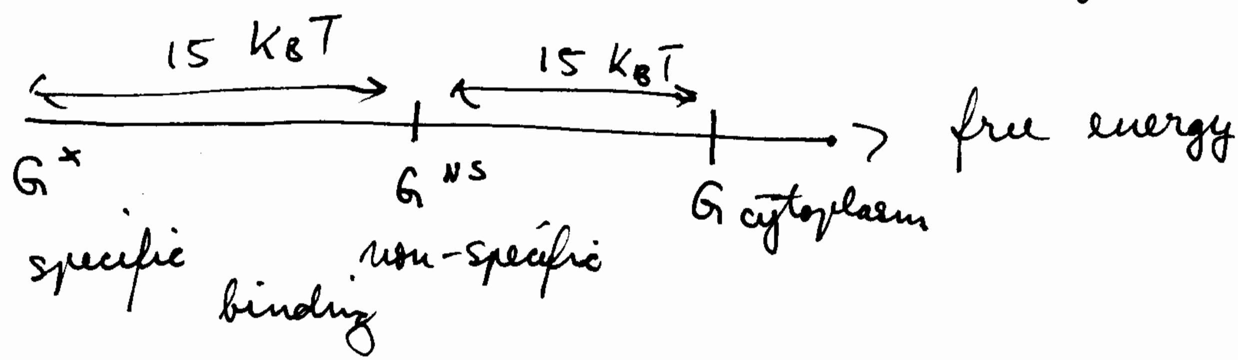
Q: Given a target site in the genome and TF conc.,
What is target occupation probability?

Consider circular genome of length $N \sim 5 \times 10^6$
$\rightarrow$ defines N possible binding sites of length L



Assume Eq. between activated TF and each S_j
i.e. fast search time (S_0 : target sequence)

$$\frac{[P][D_j]}{[P D_j]} = K_j \quad \text{(in vitro dissoc const.)}$$

$$\propto e^{+G(D_j)/kT}$$

often given as $\Delta G(D_j) = G(D_j) - G_{cyto}$

Consider $\underline{\text{one}}$ activated TF in cell

$$f(S_0) = \frac{1}{1 + \sum_{j \neq 0} K_0/K_j} = \frac{1}{1 + \sum_j e^{(G_0 - G_j)/kT}}$$

$[P]_{tot} = 1$ nm

$G_j = G(D_j)$

occupation of target site depends on
competition from rest of genome

even though $K_0 \ll K_{j \neq 0}$, N is also large

simple estimate :

$$G_0 = G^* \qquad G_{j \neq 0} = G^{ns}$$

$$\sum_j e^{(G_0 - G_j)/kT} = N e^{(G^* - G^{ns})/kT}$$

$$= N \cdot e^{-15} = 5 \times 10^6 \cdot e^{-15} \sim 1$$

comparable! coincidence?

Multiple TFs direct: tricky (fermions in heterogeneous potential w. IA)

$$f(S_0) = \frac{1}{1 + \tilde{K}_0 / [P]_{tot}}$$

$$\tilde{K}_0 = \sum_j \frac{K_0}{K_j} = \sum_j e^{(G_0 - G_j)/kT}$$

measured in unit of # / cell [nM]

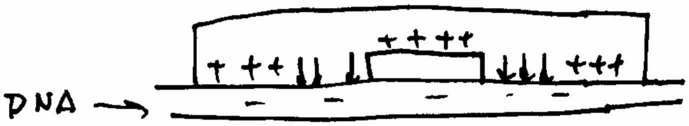
Problem :

not all G_j are known
and will not be known (impractical)

→ build model based on more limited info

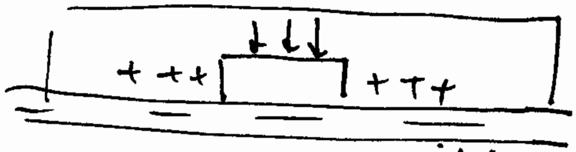
Model of TF-DNA 1A
(Winter, Berg, von Hippel '81)

- two conformations in TF-DNA binding



specific
allows AA-nucleotide pairing

$G^{sp}(s)$



non-specific
maximizes electrostatic attraction
no AA-ns pairing

G^{ns}

For any sequence, protein will bind in either conf.

$$G(s) = \min\left(G^{sp}(s), G^{ns}\right)$$

$$e^{-G(s)/kT} = e^{-G^{sp}(s)/kT} + e^{-G^{ns}/kT}$$

Exactly: Two populations, need to add Boltzmann weights of both

$$\tilde{K}_0 = \sum_{j=1}^{N} e^{-(G_j - G_0)/kT}$$

$$= N e^{-(G^{ns} - G_0)/kT} + \sum_{j=1}^{N} e^{-(G_j^{sp} - G_0)/kT}$$

let strongest binding seq be s^* with energy

$$G^{sp}(s^*) = G^*$$

additive form of less specific binding

$$G^{sp}(s) = G^* + \sum_{i=1}^{L} g_i(b_i)$$

with $g_i = 0$ if $b_i = b_i^*$

$$\tilde{K}_0 = \underbrace{\sum_j e^{-(G_j - G^*)/kT}}_{N\, e^{-(G^{us} - G^*)/kT}} \cdot \underbrace{e^{-(G^* - G_0)/kT}}_{e^{-\sum_{i=1}^{L} g_i(b_i^{(0)})}}$$

$$\sum_{j=1}^{N} e^{\pm \sum_{i=1}^{L} g_i(b_i^{(j)})/kT}$$

$\rightarrow$ can find $\tilde{K}_0$ given $G^{us} - G^*$

$g_i(b)$