Rate of RNA production = Bound RNAP \times \text{Rate of transcription when RNAP bound}

\text{Fraction of time RNAP bound}

Amount of RNA = \frac{\text{Rate of RNA production}}{\text{Rate of RNA degradation}}

Amount of protein = T \times \text{Amount of RNA}

in E. coli \( T = 10^3 \), i.e. about 1000 proteins per mRNA.

So, what we wish to know is:

How does Pbound depend on binding strength of RNAP to promoter DNA, and on number of polymerases?
Now we make the key assumption: $P_{\text{bound}}$ is determined by Boltzmann's law, or in other words, binding is in equilibrium!

(One can make various arguments why this might be true, but at the end of the day, Nature (experiment) are the final arbiter!)

1. One promoter site + one non-specific site

- $N_{\text{us}} = 2$
- $P = 1$ (number of RNA's)

\[
P_{\text{bound}} = ?
\]

**Boltzmann formula**

\[
P_{\text{bound}} = \frac{e^{-\frac{E_s}{k_B T}}}{e^{-\frac{E_n}{k_B T}}}
\]

- State: $P_{\text{bound}}$
- Probability: $e^{-\frac{E_s}{k_B T}}$
- Energy: $E_s$
- Specific factors
- $E_s < E_n$
- Non-specific factors
- Joules

**Boltzmann factors**

$k_B T$ = thermal energy @ room temperature = 300K

\[
k_B T = 1.38 \times 10^{-23} \text{ J/K} \cdot 300 \text{K} = 4.16 \times 10^{-21} \text{J}
\]

Typical energy of motion of a molecule @ temperature $T$. 

Fundamental formula of Statistical Mechanics
Example:

- solvent molecules
  - typical non-specific binding energy
    \[ E_{ns} = -2k_BT \]
  - \[ e^{-E_{ns}/k_BT} = e^2 \approx 7 \]

\[ \text{Pbound} = \frac{e^2}{e^0} \approx 7 \]  
Probability that RNAP is bound is 7 x greater than it not being bound!

General:

\[ \text{Pbound} = 1 - \text{Punbound} \]

\[ \text{Punbound} = \frac{e^{-E_{NS}/k_BT}}{e^{-E_{NS}/k_BT} + e^{-\Delta\epsilon/k_BT}} \]

\[ \text{Pbound} = e^{-\Delta\epsilon/k_BT} \]

\[ (1 + e^{-\Delta\epsilon/k_BT}) \text{Pbound} = e^{-\Delta\epsilon/k_BT} \]

\[ \text{Pbound} = \frac{e^{-\Delta\epsilon/k_BT}}{1 + e^{-\Delta\epsilon/k_BT}} \]

With the numbers in the example (\( \Delta\epsilon = -2k_BT \)):

\[ \text{Pbound} = \frac{7}{8} \]  
(\( \text{Punbound} = \frac{1}{8} \))

---

2. One promoter site + \( N_{ns} = 5 \times 10^6 \) non-specific binding sites

(Still \( P = 1 \) i.e. one polymerase)

<table>
<thead>
<tr>
<th>state</th>
<th>energy</th>
<th>multiplicity</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( E_S )</td>
<td>1</td>
<td>( 1 \times e^{-E_S/k_BT} )</td>
</tr>
<tr>
<td></td>
<td>( E_N )</td>
<td>( N_{ns} )</td>
<td>( N_{ns} \times e^{-E_{ns}/k_BT} )</td>
</tr>
</tbody>
</table>

\[ \text{Pbound} = \frac{e^{-E_S/k_BT}}{e^{-E_S/k_BT} + N_{ns} e^{-E_{ns}/k_BT}} \]
\[ P_{\text{bound}} = \frac{\frac{4}{\text{Nns}} e^{-\frac{\Delta E}{k_B T}}}{\frac{4}{\text{Nns}} e^{-\frac{\Delta E}{k_B T}} + 1} \]

3. One promoter site, \( \text{Nns} (= 5 \times 10^9) \) non-specific sites, and \( P > 1 \) (= 3000)

<table>
<thead>
<tr>
<th>State</th>
<th>Energy</th>
<th>Multiplicity</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nns</td>
<td>( E_s + (P-1)E_n )</td>
<td>( \text{Nns!} )</td>
<td>( \frac{\text{Multiplicity} \times e^{-\frac{E_n}{k_B T}}}{\text{Nns!} (\text{Nns-P}+1)!} )</td>
</tr>
<tr>
<td>( P ) polymerses</td>
<td>( P \text{E}_n )</td>
<td>( \frac{\text{Nns!}}{P! (\text{Nns-P}+1)!} )</td>
<td>( \frac{\text{Multiplicity} \times e^{-\frac{E_n}{k_B T}}}{\text{Nns} \times (\text{Nns-P})!} )</td>
</tr>
</tbody>
</table>

From the states and weights table:

\[
\frac{\text{weight of bound state}}{\text{weight of unbound state}} = \frac{\text{Nns!}}{(P!)! (\text{Nns-P+1})!} \frac{e^{-\left( E_s + (P-1)E_n \right) / k_B T}}{e^{-\frac{E_n}{k_B T}}} \]

\[
= \frac{\text{P!} (\text{Nns-P})!}{(P!)! (\text{Nns-P+1})!} \frac{e^{-\left( E_s - E_n \right) / k_B T}}{e^{-\frac{E_n}{k_B T}}} \]

\[
= \frac{\text{P!} \times (\text{Nns-P+1})!}{(P!)! (\text{Nns-P+1})!} \frac{e^{-\Delta E / k_B T}}{e^{-\Delta E / k_B T}} \]

\[
= \frac{P}{\text{Nns-P+1}} e^{-\Delta E / k_B T} \]
Since \( P \ll N_{ns} \) (\( P = \text{few thousand polymerases}, N_{ns} = \text{few } 10^6 \)) we can use a simpler formula: \( \frac{P}{N_{ns}} e^{-\Delta \Delta E/k_BT} \)

Finally, substituting the simpler formula for the ratio of weights:

\[
\text{Pbound} = \frac{P}{N_{ns}} e^{-\Delta \Delta E/k_BT} \div \frac{P}{N_{ns}} e^{-\Delta \Delta E/k_BT} + 1
\]

For \( \Delta \Delta E \approx 3k_BT \Rightarrow e^{-\Delta \Delta E/k_BT} = e^3 \approx 20 \) typical binding energy (\( \Delta \Delta E \)) for an E-coli promoter

\[
\text{Pbound} = \frac{\frac{3 \times 10^3}{5 \times 10^6} \cdot 20}{\frac{3 \times 10^3}{5 \times 10^6} \cdot 20 + 1} = \frac{10^{-2}}{10^{-2} + 1}
\]

\[
\text{Pbound} = \frac{1}{100} \text{ (fairly small)}
\]

Note that this calculation could be used for any binding problem, where \( P \) is the \# of ligands in solution (i.e. free) that can bind to receptor, \( N_{ns} \) is replaced by number of places the ligands can be in solution \( (= \frac{V}{V_w} \) where \( V \) is the volume taken up by 1 ligand and \( V_w \) the volume of the solution), and \( \Delta \Delta E \) is the ligand-receptor binding energy.

Also: \( \text{Pbound} = \frac{L/V}{(V/V_w) e^{-\Delta \Delta E/k_BT} + 1} = \frac{[L]/K_d}{[L]/K_d + 1} \)

where \( \frac{V}{V_w} \) is the \# of solution sites

\[
K_d = \frac{e^{\Delta \Delta E/k_BT}}{[L]/K_d + 1}
\]

\([L] = \text{concentration of ligand}
\]

\( K_d = \frac{e^{\Delta \Delta E/k_BT}}{[L]/K_d + 1} \) dissociation constant.
1. Cell has 5 proteins. What is the chance that daughter has all 5 (the other daughter has 0)?

\[
P_5 = \left(\frac{1}{2}\right)^5 = \frac{1}{32}
\]

\[
P_5 = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{32}
\]

2. What is the probability that daughter cell 1 gets 3 of the 5 proteins?

Each of these outcomes has probability

\[
\text{Outcome} = \frac{1}{32}
\]

\[
\text{# of desired outcomes} = 10
\]

\[
\Rightarrow P_3 = \frac{10}{32}
\]

Counting like this gets hard (or impossible!) when the number of proteins gets large. Instead use factorials (\(n! = n(n-1)(n-2)\ldots\)):

\[
\begin{array}{c}
\text{5 proteins in mother cell.} \\
\text{How many ways one there of assigning 3 to daughter 1 and 2 to daughter 2?}
\end{array}
\]
A trick that at first might seem to complicate things. Color each daughter label \((1 \text{or } 2)\) with a different color:

\[
\begin{align*}
1 & 1 1 2 2 \\
1 & 1 2 1 2 \\
1 & 1 2 2 1 \\
1 & 2 1 1 2 \\
1 & 2 1 2 1 \\
1 & 2 2 1 1
\end{align*}
\]

All 5 labels are now distinguishable so the \# of ways of assigning them to the 5 protons:

\[
5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120
\]

\# of ways of choosing which gets assigned 1
\# of ways of assigning 1 to a
etc.

This way of counting will overcount by a lot since, for example, all of these:

\[
\begin{align*}
1 & 1 1 2 2 \\
1 & 1 2 1 2 \\
1 & 1 2 2 1 \\
1 & 2 1 1 2 \\
1 & 2 1 2 1 \\
1 & 2 2 1 1
\end{align*}
\]

will be counted separately.

(If we ignore the colors on the daughter cell labels \((1 \text{and } 2)\), which we wish to do at the end of the calculation, all these assignments are one and the same!)

\[
3! = 6 \text{ permutations of the } 3 \text{ colors on } 1
\]

Therefore, we need to divide by \(3!\) and \(2!\), accounting for the permutations of the color labels. Final result: # of ways of assigning 11122 is:

\[
\frac{5!}{3!2!} = \frac{120}{12} = 10.
\]
To describe diffusion mathematically we will use the one-dimensional random walk:

\[ \begin{array}{cccccccc}
\ldots & a & \ldots & 0 & a & \ldots & 0 & a & \ldots \\
\ldots & a & \ldots & a & a & \ldots & a & a & \ldots \\
\end{array} \]

Diffusing particle starts at \( x = 0 \), and hops to the left and to the right a distance \( a \), with rate \( k \). (Assumptions of model)

What is the probability of finding particle at position \( x \) at time \( t \)?

To compute \( p(x, t) \) we need an update rule, which tells you:

- Given \( p(x, t) \), i.e., whole probability distribution at time \( t \), what is the distribution at time \( t + \Delta t \)?

\[ p(x, t + \Delta t) = ? \]

We can compute this by noticing that if the particle is at \( x \) at time \( t \), it had to be either:

1. \( x \) at time \( t + \Delta t \)
2. \( x - a \) at time \( t + \Delta t \)
3. \( x + a \) at time \( t + \Delta t \)

as long as \( \Delta t \) is tiny! \( \Delta t < k \) or \( k \Delta t < 1 \)

So:

\[ p(x, t + \Delta t) = p(1) + p(2) + p(3) \]

\[ = p(\text{particle at } x \text{ AND no hop}) + p(\text{particle at } x-a \text{ AND hop}) + p(\text{particle at } x+a \text{ AND hop}) \]

Now, rules of probability say that \( p(A \text{ and } B) = p(A) \cdot p(B) \) as long as \( A \) and \( B \) are independent. We can use this since to hop or not to hop is independent of position (it always happens with rate \( k \)).
\[
P(\text{particle ex AND no hop}) = \frac{P(x,t)}{\text{particle ex no hop}} \cdot (1 - 2k\Delta t)
\]

\[
P(\text{particle ex=0 AND hop}) = \frac{P(x=0,t)}{\text{particle ex=0 hop to right}} k\Delta t
\]

\[
P(\text{particle ex>0 AND hop}) = \frac{P(x>0,t)}{\text{particle ex>0 hop to left}} k\Delta t
\]

\[\Rightarrow P(x,t+\Delta t) = P(x,t) (1 - 2k\Delta t) + P(x=0,t) k\Delta t + P(x>0,t) k\Delta t\]

Update rule for 1D random walk (model of diffusion) with a bit more work we can turn this into a differential equation:

\[
P(x+0,t) = P(x,t) + \alpha \frac{dP}{dx} + \alpha^2 \frac{d^2P}{dx^2} + \times \text{ step at 2nd order}
\]

\[
P(x-0,t) = P(x,t) - \alpha \frac{dP}{dx} + \alpha^2 \frac{d^2P}{dx^2} + \times
\]

\[
\frac{P(x,t+\Delta t) - P(x,t)}{\Delta t} = -2kP(x,t) + k(P(x,t) + \alpha \frac{dP}{dx} + \alpha^2 \frac{d^2P}{dx^2})
\]

\[\text{Diffusion equation: } \frac{dP(x,t)}{dt} = k\alpha^2 \frac{d^2P}{dx^2} \]

\[k\alpha^2 \text{ has units of } \mu^2/\text{sec and is the diffusion constant } D. \]

Solution of the Diffusion equation: \[P(x,t) = \frac{1}{\sqrt{4\pi DT}} e^{-\frac{x^2}{4DT}} \]

is a Gaussian whose variance: \[\langle x^2 \rangle = 2D \Delta t \text{ grows linearly with time} \]

(The width of the distribution: \[\sqrt{\langle x^2 \rangle} = \sqrt{2D \Delta t} \] )
Solution in pictures:

\[ P(x) \]

\[ c = 0 \quad \text{Particle starts at } x = 0 \quad \text{therefore } P(x; t) \text{ is initially very sharply peaked at } x = 0 \]

\[ \sigma_x = \sqrt{2Dt} \]

\( \sigma_x \) tells us how far from the origin it is likely (with some confidence) to find the particle.

So for a protein in a cytoplasm: \( D = 10 \mu m^2/sec \).

Therefore, after time \( t = 5 \) seconds:

\[ \sigma_x = \sqrt{2Dt} \]

\[ \sigma_x = \sqrt{100 \mu m^2} \]

\[ \sigma_x = 10 \mu m \]

It is likely the particle will have diffused away over a distance of 10 \( \mu m \). On cellular scales diffusion is pretty fast! Unless we’re looking at e.g. long neurons...

\[ \begin{array}{c}
\text{nerve} \\
\text{cell} \\
\text{axon} \\
\text{10 cm}
\end{array} \]

\[ \sigma_x = 10 \text{ cm} = 10^5 \mu m \]

\[ 10^5 \mu m = \sqrt{2D \cdot t} \quad \Rightarrow \quad t = \frac{10^{10}}{20} \text{ sec} \]

\[ t = 5 \times 10^8 \text{ sec} \approx 10 \text{ years}!! \]

\( \Rightarrow \) Need to wait \( \sim 10 \) yrs for a protein to diffuse 10 cm.