

La vida en entornos abarrotados y desordenados



Material tomado de este libro

El interior celular está tan abarrotado que la distancia entre las proteínas vecinas es comparable al tamaño de la proteína en sí. Del mismo modo, la membrana celular contiene un gran número de proteínas de diferentes tipos sobre un fondo lipídico muy variado. Esto significa que el entorno material de las células es muy diferente de las condiciones que se encuentran en la mayoría de los experimentos bioquímicos y que aparecen en muchos de los modelos que fuimos viendo.

En particular, hemos ignorado los efectos de este hacinamiento al explotar ideas simples sobre la unión que hacen uso del límite de la "solución diluida". De manera similar, nuestro tratamiento de la cinética de difusión se ha construido alrededor de una imagen de caminata aleatoria en la que las moléculas son libres de deambular sin las trabas de interacciones con caminantes vecinos.

Acá veremos algunos ejemplos sencillos que relajan la hipótesis de la solución diluida.

Recordemos contenido de una bacteria que mide $\sim 1\mu\text{m}$ de largo

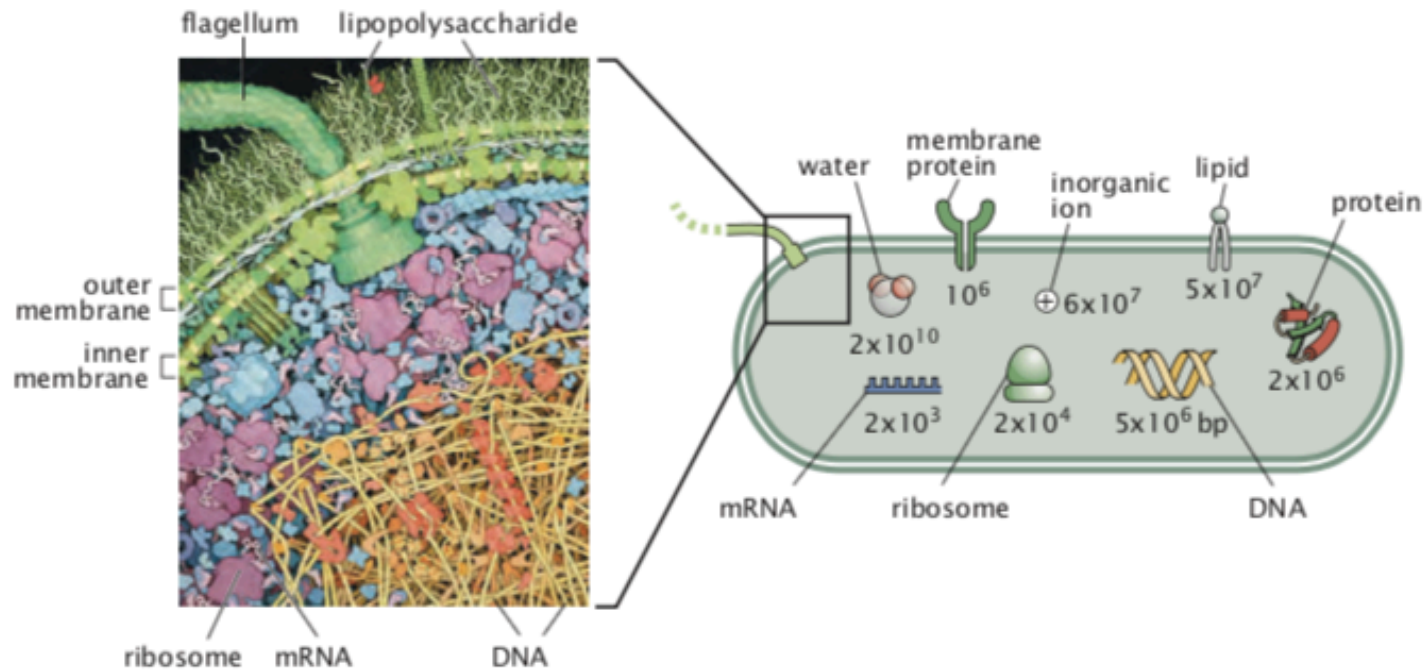


Figure 2.4: Molecular contents of the bacterium *E. coli*. The illustration on the left shows the crowded cytoplasm of the bacterial cell. The cartoon on the right shows an order-of-magnitude molecular census of the *E. coli* bacterium with the approximate number of different molecules in *E. coli*. (Illustration of the cellular interior courtesy of D. Goodsell.)

Table 2.1: Observed macromolecular census of an *E. coli* cell. (Data from F. C. Neidhardt et al., Physiology of the Bacterial Cell, Sinauer Associates, 1990 and M. Schaechter et al., Microbe, ASM Press, 2006.)

Substance	% of total dry weight	Number of molecules
Macromolecules		
Protein	55.0	2.4×10^6
RNA	20.4	
23S RNA	10.6	19,000
16S RNA	5.5	19,000
5S RNA	0.4	19,000
Transfer RNA (4S)	2.9	200,000
Messenger RNA	0.8	1,400
Phospholipid	9.1	22×10^6
Lipopolysaccharide (outer membrane)	3.4	1.2×10^6
DNA	3.1	2
Murein (cell wall)	2.5	1
Glycogen (sugar storage)	2.5	4,360
Total macromolecules	96.1	
Small molecules		
Metabolites, building blocks, etc.	2.9	
Inorganic ions	1.0	
Total small molecules	3.9	

Cálculo sencillo:

Supongamos un cubo de 1 μm de lado, supongamos que hay 10^6 moléculas de proteína. Tenemos entonces 10^6 moléculas/(1 μm^3). Esto significa que cada molécula tiene un volumen $\sim 10^{-6} \mu\text{m}^3 = 10^3 \text{ nm}^3$ a “su disposición”. O sea, la separación media entre dos moléculas es $\sim 10 \text{ nm}$: del mismo orden que el tamaño de la molécula en sí!

Muchos de los componentes responsables de la organización estructural y las propiedades mecánicas de las células y los organismos vivos tienen una estructura de fibras alargadas.

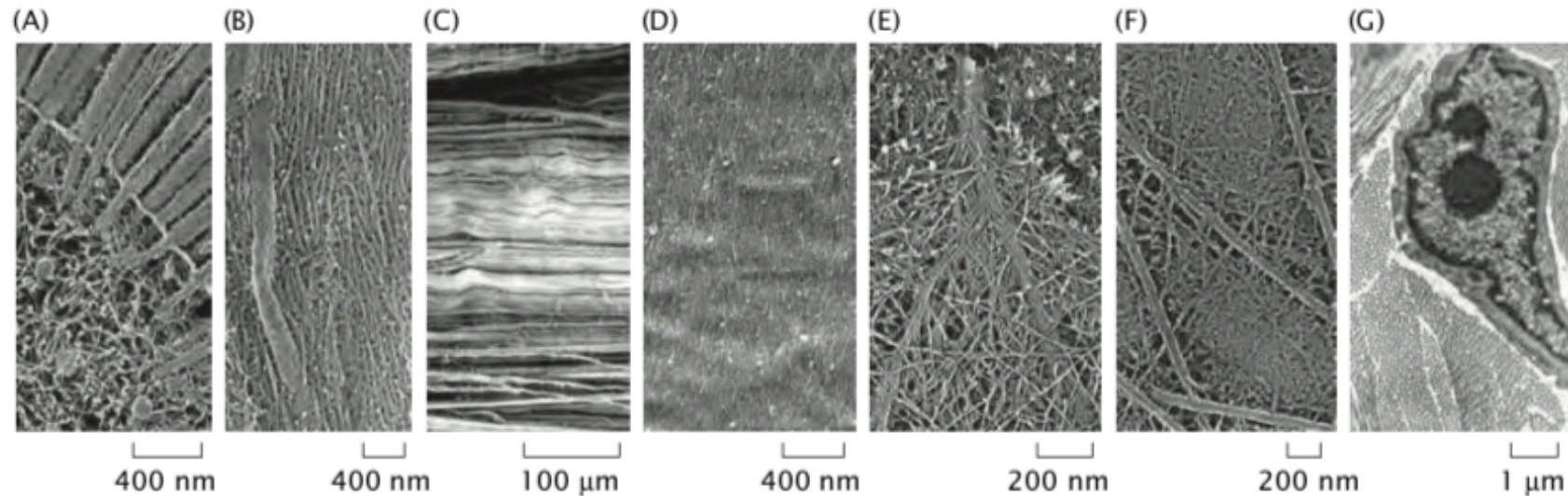


Figure 14.1: Gallery of filamentous network architectures found in cells and extracellular matrix. (A) The upper edge of an intestinal epithelial cell reveals tightly packed bundles of actin in the microvilli and a less organized but still dense actin meshwork just below the membrane. (B) Within the axon of a nerve cell, aligned and tightly packed bundles of neurofilaments and microtubules allow the axon to grow very long and serve as tracks for transporting organelles such as mitochondria. (C) In animal ligaments such as this one from a rat, aligned collagen bundles resist shear stresses. (D) The cell wall of plants is made of aligned bundles of cellulose for organisms ranging from this tiny weed *Arabidopsis* to the mighty Redwood. (E) At the leading edge of a moving cell, actin filaments in webs and bundles grow to push the membrane forward. (F) Animal tissues made of epithelial cells usually include a basement membrane lining underneath the cells to improve the mechanical strength of the tissue. The basement membrane is not a phospholipid membrane but rather a felt-like mat of collagen, proteoglycans, and other filamentous molecules. (G) Cells such as the fibroblast shown here navigate through densely packed extracellular matrix. In this thin section, collagen fibrils have been sliced in several different orientations, revealing their regular packing. (A, F, courtesy of J. Heuser; B, adapted from N. Hirokawa, *J. Cell Biol.* 94:129, 1982; C, adapted from P. P. Provenzano et al., *J. Orthop. Res.* 20:975, 2002; D, adapted from D. H. Burk and Z. H. Ye, *Plant Cell* 14:2145, 2002; E, adapted from T. M. Svitkina et al., *J. Cell Biol.* 160:409, 2003; G, adapted from P. C. Cross and K. L. Mercer, *Cell and Tissue Ultrastructure*. W. H. Freeman, 1993.)

Volvamos al “crowding”

El “crowding” afecta el equilibrio.

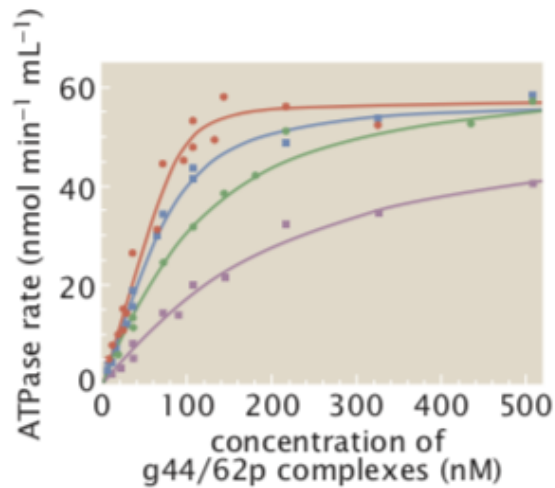


Figure 14.3: ATPase rate associated with T4 DNA replication. The different curves measure the ATPase rate as a function of the g44/62p (clamp loader) protein concentration as measured using different concentrations of polyethylene glycol as a crowding agent. The concentrations of polyethylene glycol going from the bottom to the top curve are 0, 2.5, 5, and 7.5 weight percent. (Adapted from T. C. Jarvis et al., *J. Biol. Chem.* 265:15160, 1990.)

En la Figura 14.3 se muestra un ejemplo de cómo el “binding” puede verse alterado por efectos de la aglomeración. En este caso, la medición examina los complejos proteicos accesorios asociados con la replicación del ADN en un bacteriófago. Acá se ve el efecto del crowding sobre la probabilidad de unión de los productos génicos 45 (pinza deslizante) y 44/62 (cargador de pinza) que luego se unen a la ADN polimerasa y mejoran su actividad. Estas cuatro moléculas juntas son suficientes para producir la síntesis de la cadena principal en la replicación del ADN. En la figura se muestra la tasa de hidrólisis de ATP por el cargador de pinza (g44 / 62p) para una concentración constante de g45p. Se supone que esta curva refleja el grado de unión de g45p a g44 / 62p. Como control, se demostró que cambiar la concentración de las moléculas que aumentan el crowding no cambiaba por sí mismo la actividad catalítica de g44 / 62p. Los datos presentados en esta figura no son más que un ejemplo de un fenómeno generalizado en el que la adición de moléculas inertes de apiñamiento modifica los equilibrios bioquímicos.

Y también la dinámica.

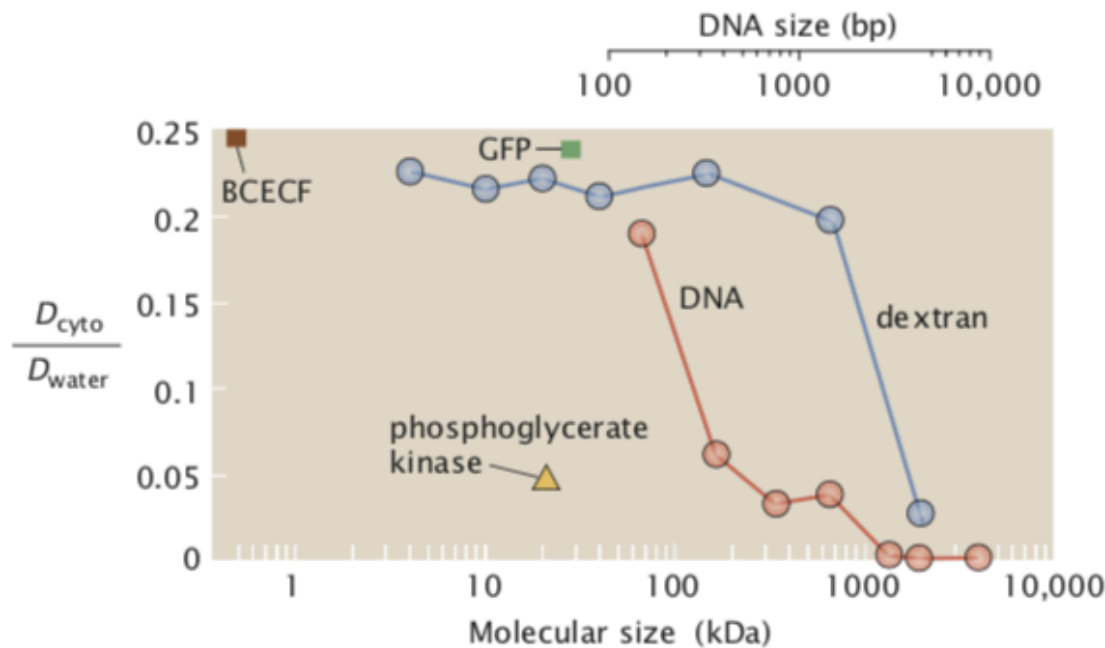


Figure 14.4: Diffusion constants in cells. The plot shows the ratio of the measured cellular diffusion constant to that for the same molecule in water for several different molecules, including a series of DNA molecules of different sizes and BCECF, a small fluorescent molecule. (Adapted from A. S. Verkman, *Trends Biochem. Sci.* 27:27, 2002.)

En la Figura se muestra el coeficiente de difusión relativo entre el valor que toma en el interior de la célula y en una solución acuosa como función del tamaño de la partícula difusora. Estas curvas muestran una disminución de cuatro veces en la constante de difusión, con un efecto aún más fuerte para moléculas más grandes.

Para poder hacer algunos cálculos sobre el efecto del crowding sobre el equilibrio, antes algo sobre la entropía y las configuraciones microscópicas

La termodinámica mira a los sistemas de muchas partículas como un todo. No se pregunta qué le pasa a cada una de ellas, sino cómo son los valores de algunas cantidades macroscópicas (promediadas sobre todas ellas).

Pero no todas las partículas hacen lo mismo, están en distintos lugares, se mueven con distintas velocidades, etc, etc, etc. La entropía es una medida de esa variabilidad alrededor del comportamiento promedio que capturan otras cantidades macroscópicas como la energía.

¿Qué “cuantifica” la entropía? Cuántos estados “microscópicos” son compatibles con el estado macroscópico que estoy viendo.

La entropía es proporcional al logaritmo del número de estados microscópicos compatibles con el estado macroscópico (que me da, por ejemplo, la energía que estoy viendo, etc).

¿Qué significa que en el equilibrio la entropía de un sistema aislado sea máxima? Significa que el estado de equilibrio es aquél al que le corresponde el mayor número posible de estados microscópicos.

Veamos un ejemplo sencillo

Supongamos N partículas que se reparten entre dos volúmenes, V_A y V_B .

La probabilidad de que una partícula “caiga” en V_A es V_A/V .

La probabilidad de tener N_A moléculas en V_A está dada por la distribución binomial, $\text{Bin}(V_A/V, N)$.



$$\begin{aligned} P(N_A|N) &= \frac{N!}{N_A!(N - N_A)!} \left(\frac{V_A}{V}\right)^{N_A} \left(1 - \left(\frac{V_A}{V}\right)\right)^{N - N_A} \\ &= \binom{N}{N_A} \left(\frac{V_A}{V}\right)^{N_A} \left(1 - \left(\frac{V_A}{V}\right)\right)^{N - N_A} \end{aligned}$$

$$P(N_A, N_B) = \frac{N!}{N_A!N_B!} \left(\frac{V_A}{V}\right)^{N_A} \left(\frac{V_B}{V}\right)^{N_B}$$

Algunas propiedades de la binomial

Aproximación de Stirling para números grandes: $\ln N! \approx N \ln N - N + \frac{1}{2} \ln \left[\left(2N + \frac{1}{3} \right) \pi \right]$

(Significa que para N suficientemente grande es: $N! \sim N^N$)

Aproximación de Stirling para distribución binomial, $\text{Bin}(p, N)$ (n y N grandes,):

$$P(n|N) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}$$

$$\ln P(n|N) \approx N \ln N - n \ln n - (N-n) \ln(N-n) \\ + n \ln p + (N-n) \ln(1-p)$$

¿Para qué valor de n es $P(n|N)$ máximo? $\frac{\partial}{\partial n} \ln P(n|N) = 0$

Es máximo para n_o tal que: $\frac{n_o}{N - n_o} = \frac{p}{1-p}$ o sea: $n_o = pN$

Pero: $\langle n \rangle = pN$, o sea, tiene el máximo en el valor medio.

Más aún, para N suficientemente grande (pN finito) la $\text{Bin}(p, N)$ se puede aproximar por:

$$P(n|N) \approx \frac{1}{\sqrt{2\pi p(1-p)N}} \exp \left[-\frac{(n - pN)^2}{2p(1-p)N} \right]$$

Gaussiana con igual valor medio y varianza que la Binomial

Volvamos a los volúmenes y supongamos N_A y N grandes

Antes

$$\langle N_A \rangle = N \left(\frac{V_A}{V} \right) \quad \langle N_B \rangle = N \left(\frac{V_B}{V} \right) \quad \frac{\langle N_A \rangle}{V_A} = \frac{\langle N_B \rangle}{V_B} = \frac{N}{V}$$

Concentración uniforme!

Sabemos que el máximo de la distribución ocurre para $N_A = \langle N_A \rangle$.

O sea, es más probable que las partículas se distribuyan de un modo uniforme. Eso significa que hay más “configuraciones” microscópicas que me dan esa distribución uniforme que otras distribuciones.

Por otro lado, desvío standard de N_A :

$$\begin{aligned} \delta N_A &= \left[N \left(\frac{V_A}{V} \right) \left(1 - \frac{V_A}{V} \right) \right]^{1/2} & \frac{\delta N_A}{\langle N_A \rangle} &= \frac{1}{\langle N_A \rangle} \left[\langle N_A \rangle \left(\frac{V_B}{V} \right) \right]^{1/2} = \sqrt{\frac{1}{\langle N_A \rangle}} \sqrt{\frac{V_B}{V}} \\ &= \left[N \left(\frac{V_A}{V} \right) \left(\frac{V_B}{V} \right) \right]^{1/2} \\ &= \left[\langle N_A \rangle \left(\frac{V_B}{V} \right) \right]^{1/2} \end{aligned}$$

que el ancho sea pequeño comparado con $\langle N_A \rangle$ si $\langle N_A \rangle$ es suficientemente grande nos dice que casi siempre estamos en el caso más probable!

¿Cómo comparo esto con las “configuraciones”?

Supongamos que tengo un solo volumen, V , y N partículas en él. Para contar “configuraciones” (asociadas solo a la posición) tengo que subdividir a V en pequeños volúmenes, v . Supongamos que $V/v \gg N$ (situación diluida).

Tomo una partícula, tengo V/v volúmenes pequeños donde ponerla. Para la segunda tengo $(V/v-1)$ y así hasta la última, $(V/v-N+1)$. Entonces, el número de configuraciones sería:

$$V/v (V/v-1) (V/v-2) \dots (V/v-N+1) \approx (V/v)^N$$

Pero si no distingo a las partículas entre sí (es decir, lo único que me interesa es qué volúmenes están ocupados), en cada una de estas configuraciones puedo intercambiar (reordenar) las N partículas. El número queda entonces:

$$(V/v)^N / N!$$

Justamente, $P(N_A, N_B) = \frac{N!}{N_A! N_B!} \left(\frac{V_A}{V}\right)^{N_A} \left(\frac{V_B}{V}\right)^{N_B}$ se puede reescribir como

$$P(N_A, N_B) = \frac{\Omega_q(N_A, V_A) \Omega_q(N_B, V_B)}{\Omega_q(N, V)} \quad \text{con} \quad \Omega_q(N, V) = \frac{V^N}{N!}$$

O sea, como el producto del número de configuraciones con N_A en V_A por el de config con N_B en V_B dividido por el número total de configuraciones para las N partículas en V .

Definiendo

$$S_q(N, V) \equiv k \ln \left(\frac{V^N}{N!} \right) + kXN$$
$$\equiv k \ln \Omega_q(N, V) + kXN$$

con k y X algunas constantes, y,
para el sistema con dos volúmenes:

$$S_{q,tot}(N_A, V_A, N_B, V_B) = k \ln [P(N_A, N_B)] + S_q(N, V)$$
$$= S_q(N_A, V_A) + S_q(N_B, V_B)$$

Notar que, dado N , $S_{q,tot}$ es máxima
para el valor, N_A , que maximiza P ,
 $N_A = \langle N_A \rangle$

Aproximando P por su valor en el máximo (dado que es una Gaussiana
con ancho relativamente pequeño) es:

$$\ln P(N_A, N_B)|_{equil} \approx -\frac{1}{2} \ln (2\pi \langle N_A \rangle (V_B/V))$$

Pero $S_q(N, V) \sim N$, por lo que $\ln(P)$ es despreciable frente a $S_q(N, V)$:

$$S_{q,tot}(N_A, V_A, N_B, V_B) = S_q(N_A, V_A) + S_q(N_B, V_B) = S_q(N, V)$$

$S_{q,tot}$ es la entropía del sistema, es igual a la suma de las entropías de
los subsistemas A y B y tiene un máximo en $N_A = \langle N_A \rangle$, $k = k_B$ ($X=0$).

Volvamos a una situación de dos volúmenes, pero con uno mucho más grande que el otro (un reservorio térmico).

Pensemos en la entropía en términos de las energías del sistema en el volumen de interés (el más pequeño), E , y en el reservorio, E_R . La energía total es constante:

$$E_T = E + E_R \qquad E_R \gg E \qquad E_T \gg E.$$

Y, como en el caso de los dos volúmenes, es: la probabilidad de que el sistema de interés tenga energía E .

$$P(E) = \frac{\Omega(E)\Omega_R(E_T - E)}{\Omega_T(E_T)}$$

Algunos cálculos:

$$\ln P(E) = \ln \Omega(E) + \ln \Omega_R(E_T - E) - \ln \Omega_T(E_T)$$

$$\ln P(E) = \ln \Omega(E) + \ln \Omega_R(E_T) - E \left(\frac{\partial \ln \Omega_R(E_T)}{\partial E_T} \right) - \ln \Omega_T + \mathcal{O}((E/E_T)^2)$$

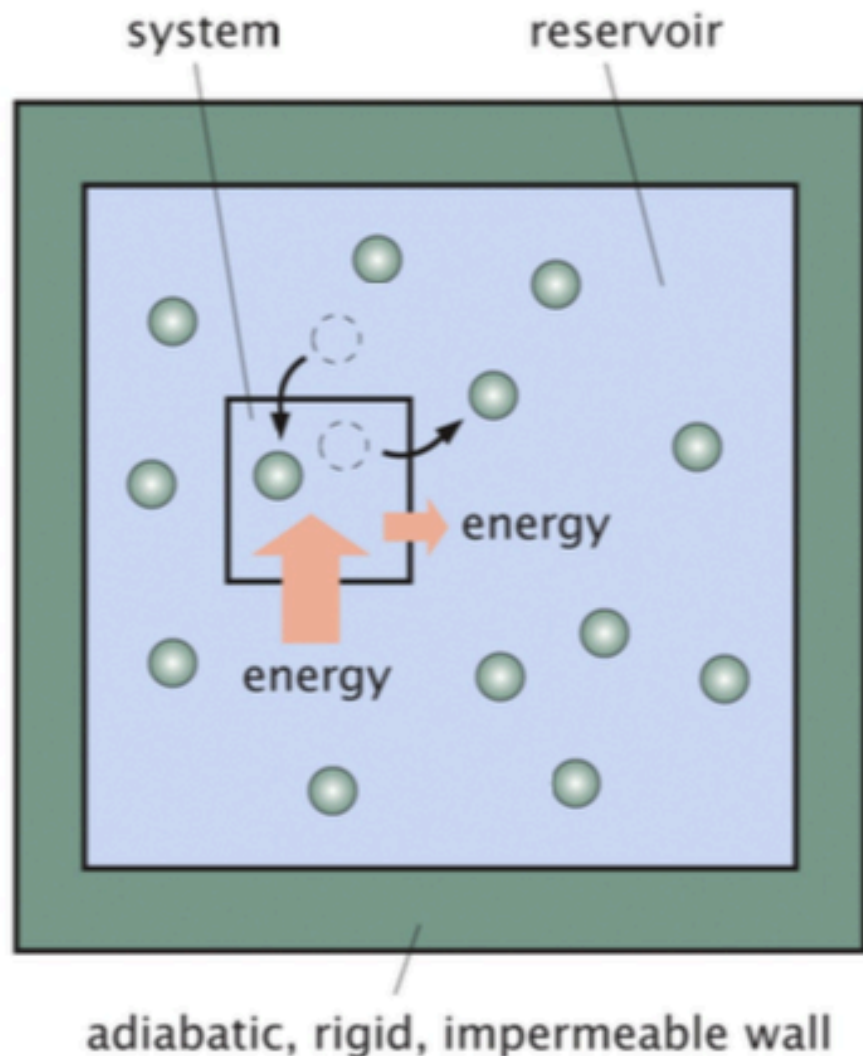
$$\ln P(E) = \ln \Omega(E) - \beta E - \ln Z \qquad P(E) = \frac{1}{Z} \Omega(E) \exp(-\beta E)$$

Relación con la termodinámica (equilibrio!), $\beta = 1/k_B T$ con T la temperatura. Z es un factor de normalización (para que $P(E)$ integrada o sumada sobre todos los valores posibles de E sea 1). Z es la función de partición.

Análogamente, se puede hacer el cálculo pensando en que el volumen no solo intercambia energía con el reservorio sino también partículas:



Grand canonical distribution



- Gibbs derived statistical model appropriate to describe a system embedded in a reservoir of energy and particles.
- Thermal and chemical equilibrium.
- Probability of system in microstate i with energy E_i and N_i particles:

$$p(E_i, N_i) = \frac{e^{-\beta(E_i - \mu N_i)}}{Z}$$

with grand partition function:

$$Z = \sum_i e^{-\beta(E_i - \mu N_i)}$$

Figure 7.9 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Acá está suponiendo un conjunto discreto de valores posibles de la energía



Grand canonical distribution

Gibb's distribution: key equations

$$p(E_i) = \frac{1}{Z} e^{-\beta(E_i - \mu N_i)}$$

$$Z = \sum_i e^{-\beta(E_i - \mu N_i)}$$

$$\langle E \rangle = \frac{1}{Z} \sum_i E_i e^{-\beta(E_i - \mu N_i)} = -\frac{\partial}{\partial \beta} \ln Z$$

$$\langle N \rangle = \frac{1}{Z} \sum_i N_i e^{-\beta(E_i - \mu N_i)} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln Z$$

...



Ligand-receptor binding III

- Consider two-states model of ligand-receptor binding and apply grand canonical distribution.
- Receptor is system, exchanging particles with reservoir.
 - Variable σ defines binding state:
 - $\sigma = 0 \rightarrow$ unbound receptor $\rightarrow E(\sigma) = 0$.
 - $\sigma = 1 \rightarrow$ bound receptor $\rightarrow E(\sigma) = \sigma\epsilon_b \leq 0$.
- Partition function and average number of bound particles given by:

$$Z = \sum_{\sigma=0}^1 e^{-\beta(\epsilon_b\sigma - \mu\sigma)} = 1 + e^{-\beta(\epsilon_b - \mu)}$$

$$\langle N \rangle = \frac{e^{-\beta(\epsilon_b - \mu)}}{1 + e^{-\beta(\epsilon_b - \mu)}}$$



Ligand-receptor binding III

STATE

WEIGHT



$\sigma = 0$

1



$\sigma = 1$

$e^{-\beta(\epsilon_b - \mu)}$

- In case of single receptor:

$$\langle N \rangle = p_{bound}$$

- Considering dilute solutions:

$$\mu = \mu_0 + k_B T \ln \frac{c}{c_0}$$

- Defining energy difference $\Delta\epsilon = \epsilon_b - \mu_0$ between taking ligand from solution and putting it on receptor gives:

$$\langle N \rangle = \frac{(c/c_0) e^{-\beta\Delta\epsilon}}{1 + (c/c_0) e^{-\beta\Delta\epsilon}}$$

Figure 7.10 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Esta última expresión se puede obtener de una manera más intuitiva con un modelo de “lattice”

Supongamos L moléculas de ligando en un lattice con Ω celdas donde meterlas. Supongamos que la energía de cada molécula es ϵ_{sol} en la solución y ϵ_{b} cuando está pegada al receptor. Hay un solo receptor, así que tenemos situaciones con L ligandos en y con $L-1$. Si $\Omega \gg L$, resulta:

Figure 6.4: States and weights diagram for ligand–receptor binding. The cartoons show a lattice model of the solution for the case in which there are L ligands. In (A), the receptor is unoccupied. In (B), the receptor is occupied by a ligand and the remaining $L - 1$ ligands are free in solution. A given state has a weight dictated by its Boltzmann factor. The multiplicity refers to the number of different microstates that share that same Boltzmann factor (for example, all of the states with no ligand bound to the receptor have the same Boltzmann factor).

STATE	ENERGY	MULTIPLICITY	WEIGHT
(A) 	$L\epsilon_{\text{sol}}$	$\frac{\Omega!}{L!(\Omega-L)!} \approx \frac{\Omega^L}{L!}$	$\frac{\Omega^L}{L!} e^{-\beta L\epsilon_{\text{sol}}}$
(B) 	$(L-1)\epsilon_{\text{sol}} + \epsilon_{\text{b}}$	$\frac{\Omega!}{(L-1)!(\Omega-L+1)!} \approx \frac{\Omega^{L-1}}{(L-1)!}$	$\frac{\Omega^{L-1}}{(L-1)!} e^{-\beta[(L-1)\epsilon_{\text{sol}} + \epsilon_{\text{b}}]}$

$$\text{weight when receptor occupied} = \underbrace{e^{-\beta\epsilon_{\text{b}}}}_{\text{bound ligand}} \times \underbrace{\sum_{\text{solution}} e^{-\beta(L-1)\epsilon_{\text{sol}}}}_{\text{free ligands}}, \quad \sum_{\text{solution}} e^{-\beta(L-1)\epsilon_{\text{sol}}} = \frac{\Omega!}{(L-1)![\Omega - (L-1)]!} e^{-\beta(L-1)\epsilon_{\text{sol}}}.$$

Análogamente

$$\text{weight when receptor occupied} = \underbrace{e^{-\beta \varepsilon_b}}_{\text{bound ligand}} \times \underbrace{\sum_{\text{solution}} e^{-\beta(L-1)\varepsilon_{\text{sol}}}}_{\text{free ligands}},$$

La función de partición resulta:

$$Z(L, \Omega) = e^{-\beta L \varepsilon_{\text{sol}}} \frac{\Omega!}{L!(\Omega - L)!} + e^{-\beta \varepsilon_b} e^{-\beta(L-1)\varepsilon_{\text{sol}}} \frac{\Omega!}{(L-1)![\Omega - (L-1)]!}.$$

Como para $\Omega \gg L$, es: $\frac{\Omega!}{(\Omega - L)!} \approx \Omega^L$, entonces:

$$p_{\text{bound}} = \frac{e^{-\beta \varepsilon_b} \frac{\Omega^{L-1}}{(L-1)!} e^{-\beta(L-1)\varepsilon_{\text{sol}}}}{\frac{\Omega^L}{L!} e^{-\beta L \varepsilon_{\text{sol}}} + e^{-\beta \varepsilon_b} \frac{\Omega^{L-1}}{(L-1)!} e^{-\beta(L-1)\varepsilon_{\text{sol}}}}.$$

Que se puede re-escribir como:

$$p_{\text{bound}} = \frac{(L/\Omega)e^{-\beta \Delta \varepsilon}}{1 + (L/\Omega)e^{-\beta \Delta \varepsilon}},$$

con

$$\Delta \varepsilon = \varepsilon_b - \varepsilon_{\text{sol}}.$$

Introduciendo: $c = L/(\Omega V_{\text{box}})$ y $c_0 = 1/V_{\text{box}}$ con V_{box} el volumen de cada celda, se obtiene

$$p_{\text{bound}} = \frac{(c/c_0)e^{-\beta \Delta \varepsilon}}{1 + (c/c_0)e^{-\beta \Delta \varepsilon}}.$$

Antes de pasar a ver cómo se modifica el cálculo en el caso de un mundo abarrotado, hagamos 2 cosas, la primera, en el pizarrón.

Acá, la segunda: aplicar estas ideas a la transcripción.

Para que la transcripción pueda proceder es necesario que la ARN polimerasa encuentre el sitio de inicio, se pegue y comience a trabajar.

Vamos a aplicar las ideas desarrolladas al caso de bacterias. Recordemos algunas cosas

En las bacterias, es una subunidad de la ARN polimerasa, el factor sigma (σ), el principal responsable de reconocer la secuencia promotora en el ADN.

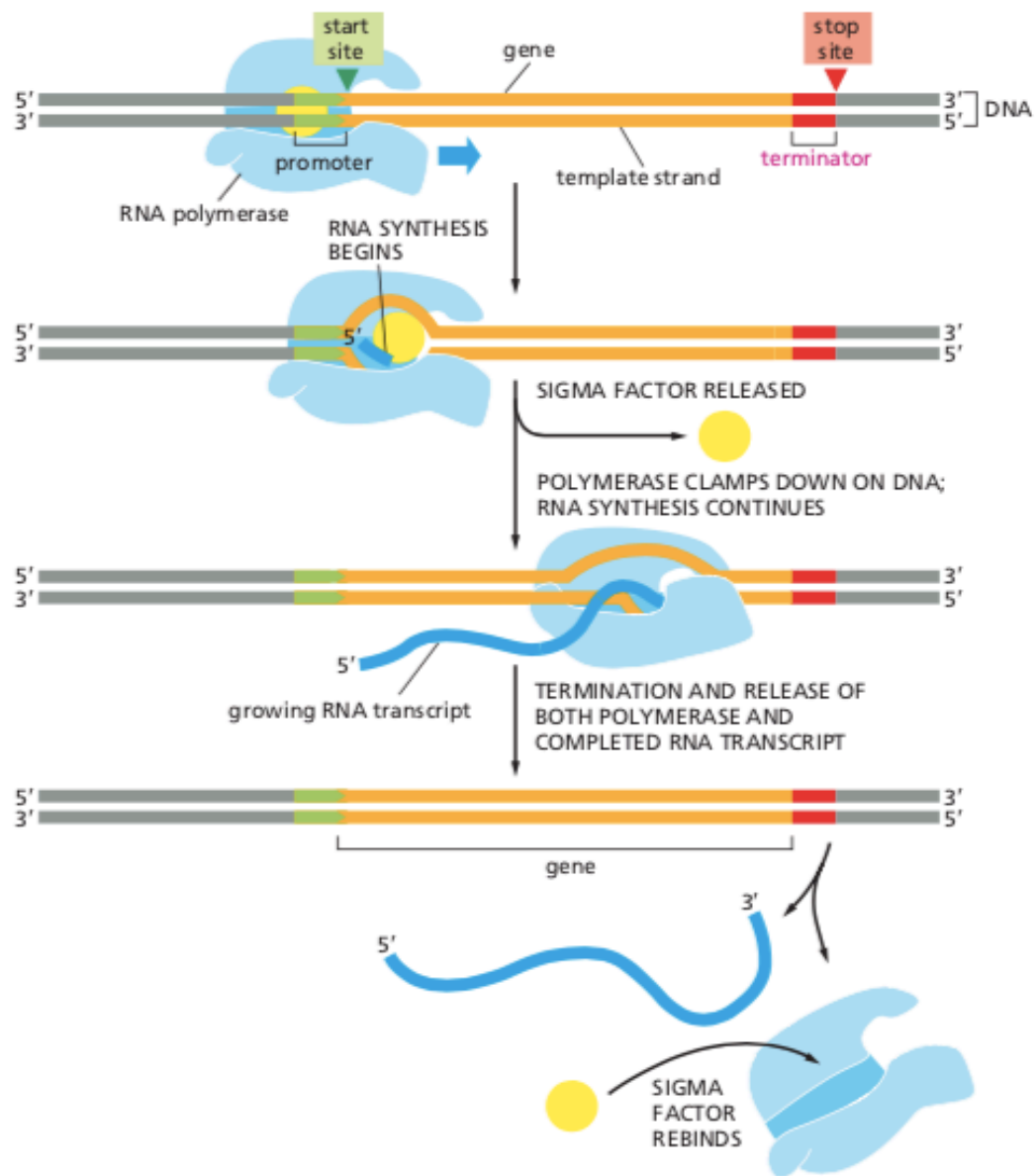


Figure 7-9 Signals in the nucleotide sequence of a gene tell bacterial RNA polymerase where to start and stop transcription. Bacterial RNA polymerase (light blue) contains a subunit called sigma factor (yellow) that recognizes the promoter of a gene (green). Once transcription has begun, sigma factor is released, and the polymerase moves forward and continues synthesizing the RNA. Chain elongation continues until the polymerase encounters a sequence in the gene called the terminator (red). There the enzyme halts and releases both the DNA template and the newly made RNA transcript. The polymerase then reassociates with a free sigma factor and searches for another promoter to begin the process again.

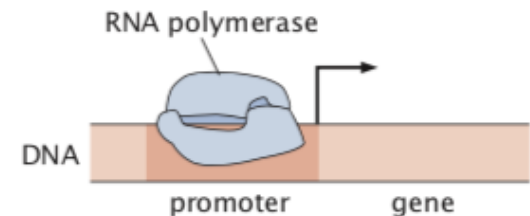


Figure 6.8: Geography of the promoter region. The promoter is the site where the RNA polymerase binds just before the transcription start site (denoted by the arrow) for the gene of interest.

El asunto es que hay muchos sitios a lo largo del ADN a los que puede pegarse la ARN polimerasa.

Por otro lado, muchas de estas moléculas de polimerasa pueden estar difundiendo (en solución)

Varios experimentos mostraron, sin embargo, que, en cada instante, la mayoría de las ~1000 moléculas de ARN polimerasas que hay en una célula de E. Coli están pegadas al ADN.

Vamos a aplicar las ideas desarrolladas al caso de bacterias. Recordemos algunas cosas

Vamos a hacer el cálculo, entonces, como si estuvieran todas pegadas y examinar cómo es la competencia entre el sitio, sobre el ADN, del promotor y los sitios no específicos donde la ARN polimerasa se puede pegar.

Supongamos que hay P moléculas pegadas en sitios no específicos.

Para calcular la función de partición y las probabilidades correspondientes pensamos al ADN subdividido en N_{NS} “cachitos”.

Llamemos ϵ_{pd}^{NS} a la energía del estado que corresponde a una molécula de RNA polimerasa pegada a un sitio no específico. Análogamente llamamos ϵ_{pd}^S a la energía cuando está pegada al promotor de interés.

Tenemos que considerar dos tipos de configuraciones: aquéllas en las que las P ARN polimerasas están pegadas a sitios no específicos y aquéllas en las que una está pegada al sitio de interés y las $P-1$ restantes no.

Pasemos al pizarrón

La parte de la función de partición que corresponde a las moléculas pegadas a los sitios no específicos es:

$$Z_{NS}(P, N_{NS}) = \underbrace{\frac{N_{NS}!}{P!(N_{NS} - P)!}}_{\text{multiplicity}} \times \underbrace{e^{-\beta P \epsilon_{pd}^{NS}}}_{\text{Boltzmann factor}}.$$

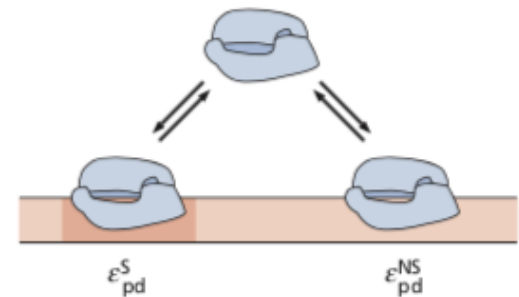
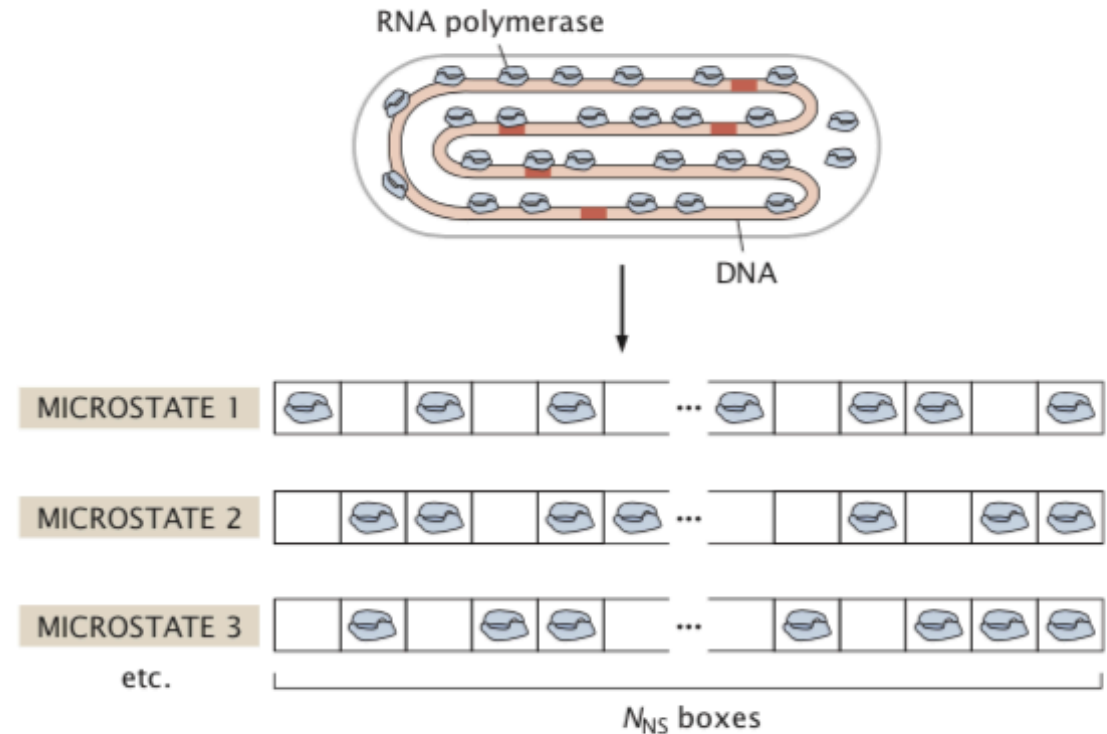


Figure 6.10: Polymerase binding energies. Illustration of the difference in binding energy for RNA polymerase when it is bound specifically (ϵ_{pd}^S) and nonspecifically (ϵ_{pd}^{NS}).

Ésta es la visión general:

Figure 6.9: The RNA polymerase reservoir. Schematic of the ways in which RNA polymerase can be distributed throughout a bacterial cell. These molecules can be bound at any of the N_{NS} nonspecific sites on the DNA, at promoters, or be distributed throughout the cytoplasm. Each microstate corresponds to a different arrangement of the polymerase molecules on these nonspecific binding sites. We consider a model in which RNA polymerase is bound exclusively on the DNA, since the cytoplasmic contribution is negligible.



Ahora calculemos la otra parte. Si llamamos: la otra parte es:

$$Z_{NS}(P, N_{NS}) = \underbrace{\frac{N_{NS}!}{P!(N_{NS} - P)!}}_{\text{multiplicity}} \times \underbrace{e^{-\beta P \epsilon_{pd}^{NS}}}_{\text{Boltzmann factor}}.$$

$$\underbrace{Z_{NS}(P - 1, N_{NS}) e^{-\beta \epsilon_{pd}^S}}_{\text{one RNAP on promoter}}.$$

En decir, Z es:

$$Z(P, N_{NS}) = \underbrace{Z_{NS}(P, N_{NS})}_{\text{no RNAP on promoter}} + \underbrace{Z_{NS}(P-1, N_{NS})e^{-\beta \epsilon_{pd}^S}}_{\text{one RNAP on promoter}}.$$

Y sus dos términos son:

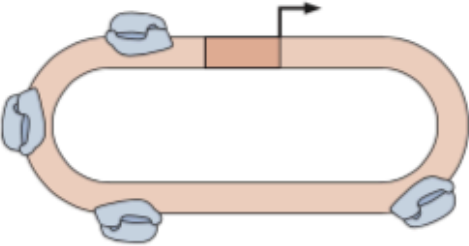
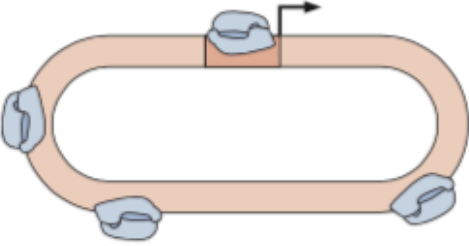
STATE	ENERGY	MULTIPLICITY	WEIGHT (MULTIPLICITY x BOLTZMANN WEIGHT)
	$P \epsilon_{pd}^{NS}$	$\frac{N_{NS}!}{P! (N_{NS}-P)!} \approx \frac{(N_{NS})^P}{P!}$	$\frac{(N_{NS})^P}{P!} e^{-P \epsilon_{pd}^{NS} / k_B T}$
	$(P-1) \epsilon_{pd}^{NS} + \epsilon_{pd}^S$	$\frac{N_{NS}!}{(P-1)! [N_{NS}-(P-1)]!} \approx \frac{(N_{NS})^{P-1}}{(P-1)!}$	$\frac{(N_{NS})^{P-1}}{(P-1)!} e^{-(P-1) \epsilon_{pd}^{NS} / k_B T} e^{-\epsilon_{pd}^S / k_B T}$

Figure 6.11: States and weights for promoter occupancy. The states are divided into two categories: those in which one of the RNA polymerase molecules is on the promoter and those for which the promoter is unoccupied.

Donde se muestra también como se pueden reescribir si $N_{NS} \gg P$, de modo que:

$$N_{NS}! / (N_{NS} - \tilde{P})! \approx (N_{NS})^P,$$

Dado que Z es:

$$Z(P, N_{\text{NS}}) = \underbrace{Z_{\text{NS}}(P, N_{\text{NS}})}_{\text{no RNAP on promoter}} + \underbrace{Z_{\text{NS}}(P-1, N_{\text{NS}})e^{-\beta\epsilon_{\text{pd}}^S}}_{\text{one RNAP on promoter}}.$$

la probabilidad de que el promotor esté pegado es:

$$p_{\text{bound}} = \frac{\frac{N_{\text{NS}}!}{(P-1)![N_{\text{NS}}-(P-1)]!}e^{-\beta\epsilon_{\text{pd}}^S}e^{-\beta(P-1)\epsilon_{\text{pd}}^{\text{NS}}}}{\frac{N_{\text{NS}}!}{P!(N_{\text{NS}}-P)!}e^{-\beta P\epsilon_{\text{pd}}^{\text{NS}}} + \frac{N_{\text{NS}}!}{(P-1)![N_{\text{NS}}-(P-1)]!}e^{-\beta\epsilon_{\text{pd}}^S}e^{-\beta(P-1)\epsilon_{\text{pd}}^{\text{NS}}}}. \quad (6.22)$$

Que, suponiendo que $N_{\text{NS}} \gg P$, de modo que: $N_{\text{NS}}!/(N_{\text{NS}}-P)! \approx (N_{\text{NS}})^P$, se reescribe como:

$$p_{\text{bound}} = \frac{\frac{P}{N_{\text{NS}}}e^{-\beta\Delta\epsilon_{\text{pd}}}}{1 + \frac{P}{N_{\text{NS}}}e^{-\beta\Delta\epsilon_{\text{pd}}}} = \frac{1}{1 + \frac{N_{\text{NS}}}{P}e^{\beta\Delta\epsilon_{\text{pd}}}}, \quad (6.23)$$

donde: $\Delta\epsilon_{\text{pd}} = \epsilon_{\text{pd}}^S - \epsilon_{\text{pd}}^{\text{NS}}.$

Figure 6.12: Promoter occupancy as a ratio. The probability of promoter occupancy by RNA polymerase is the ratio of the weights of all states with the promoter occupied to the weights of all possible states of the polymerase on the DNA.

$$p_{\text{bound}} = \frac{\sum_{\text{states}} \left(\text{Diagram of polymerase bound to promoter} \right)}{\sum_{\text{states}} \left(\text{Diagram of polymerase bound to DNA} \right) + \sum_{\text{states}} \left(\text{Diagram of polymerase bound to promoter} \right)}$$

The diagrams illustrate the states of RNA polymerase (orange) on DNA (green). The top diagram shows the polymerase bound to the promoter region, with a red arrow indicating the start of transcription. The bottom two diagrams show the polymerase bound to different regions of the DNA, with red arrows indicating the start of transcription.

$$p_{\text{bound}} = \frac{\frac{N_{\text{NS}}!}{(P-1)![N_{\text{NS}} - (P-1)]!} e^{-\beta \epsilon_{\text{pd}}^S} e^{-\beta(P-1)\epsilon_{\text{pd}}^{\text{NS}}}}{\frac{N_{\text{NS}}!}{P!(N_{\text{NS}} - P)!} e^{-\beta P \epsilon_{\text{pd}}^{\text{NS}}} + \frac{N_{\text{NS}}!}{(P-1)![N_{\text{NS}} - (P-1)]!} e^{-\beta \epsilon_{\text{pd}}^S} e^{-\beta(P-1)\epsilon_{\text{pd}}^{\text{NS}}}}}$$

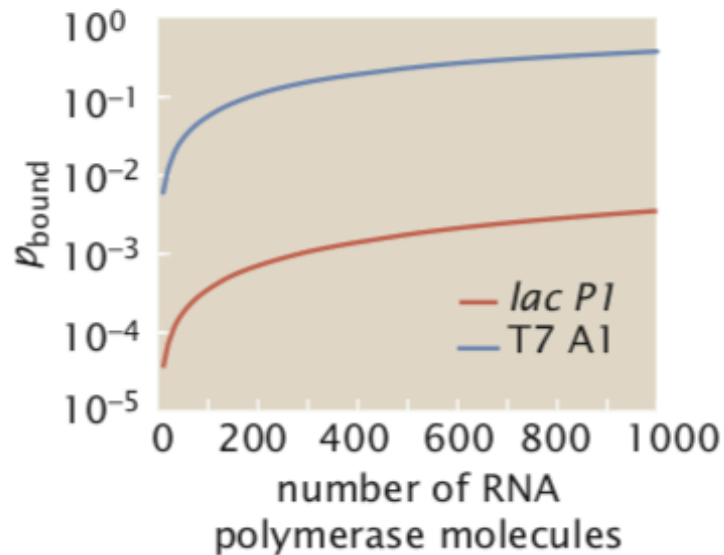


Figure 6.13: Probability of promoter occupancy as a function of the number of RNA polymerase molecules. p_{bound} is computed using values for the specific and nonspecific binding of RNA polymerase obtained *in vitro* and corresponding to the *lac* promoter, and the A1 promoter from the phage T7.

La figura muestra la probabilidad de que la ARN polimerasa se una a su promotor (en este modelo simple) en función del número de moléculas de ARN polimerasa en nuestra hipotética célula bacteriana para dos elecciones diferentes de “fuerza” del promotor: un promotor en *E. coli* para el que estimamos $\varepsilon_{\text{pd}} = -2.9 \text{ kB T}$ y el segundo caso corresponde al promotor asociado al bacteriófago T7 con una fuerza promotora caracterizada por $\varepsilon_{\text{pd}} = -8.1 \text{ kB T}$. Ambas estimaciones se pueden realizar basándose en mediciones *in vitro*.

Ahora sí, repitamos los cálculos de números de configuraciones, etc, en un caso “abarrotado”.

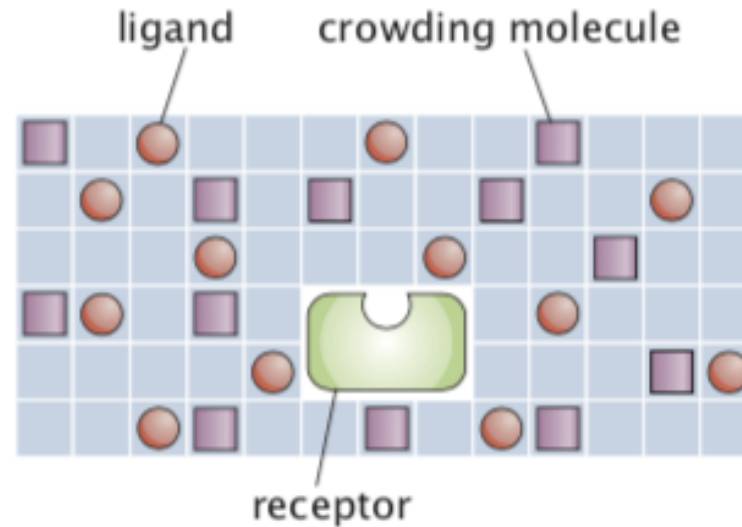


Figure 14.6: Lattice model of crowding and its effect on ligand–protein binding. The reaction volume is occupied by solvent molecules, crowding molecules, and ligands. A single ligand can bind to the protein.

Para simplificar las cosas, suponemos que el volumen de reacción se divide en Ω celdas de volumen v como en la figura. Cada uno de estos volúmenes está vacío (en realidad, ocupado por una molécula de disolvente), ocupado por un ligando u ocupado por una de las moléculas que generan el aglomeramiento. Por otro lado, hay solo un receptor al que se puede ligar el ligando.

El número total de ligandos y moléculas de “crowding” en el volumen de reacción Ω son L y C , respectivamente, como se muestra en la figura 14.6. Si consideramos la situación en la que en la solución solo hay ligandos y moléculas de crowding, entonces la función de partición sería:

$$Z_{\text{sol}}(L, C) = \frac{\Omega!}{L!C!(\Omega - L - C)!} e^{-\beta L \varepsilon_L^{\text{sol}}} e^{-\beta C \varepsilon_C^{\text{sol}}},$$

Donde $\varepsilon_L^{\text{sol}}$ y $\varepsilon_C^{\text{sol}}$ son las energías de los ligandos y de las moléculas de crowding en solución.

Acá estamos contando el número de configuraciones con L ligandos y C moléculas de crowding dado que tenemos Ω lugares donde ubicarlas y, por otro lado, estamos multiplicando esta multiplicidad por el factor de Boltzmann.

Ahora pensemos en el sitio de ligadura, que puede estar en dos estados: libre o ligado a un ligando. Los pesos de estos dos estados son:

$$Z_{\text{sol}}(L, C) \qquad Z_{\text{sol}}(L - 1, C) e^{-\beta \varepsilon_L^{\text{b}}}$$

Donde ε_L^{b} es la energía del estado que corresponde al ligando ligado al receptor.

La probabilidad de que el receptor esté ligado es entonces:

$$p_{\text{bound}} = \frac{Z_{\text{sol}}(L - 1, C) e^{-\beta \varepsilon_L^{\text{b}}}}{Z_{\text{sol}}(L - 1, C) e^{-\beta \varepsilon_L^{\text{b}}} + Z_{\text{sol}}(L, C)}.$$

Donde $\Delta \varepsilon_L \equiv \varepsilon_L^{\text{b}} - \varepsilon_L^{\text{sol}}$

Suponiendo que: $\Omega - L - C \gg 1$ la probabilidad se puede reescribir como:

$$p_{\text{bound}} = \frac{1}{1 + \frac{\Omega - L - C}{L} e^{\beta \Delta \varepsilon_L}},$$

Comparemos este resultado con el que habíamos obtenido antes:

$$p_{\text{bound}} = \frac{(L/\Omega) e^{-\beta \Delta \varepsilon}}{1 + (L/\Omega) e^{-\beta \Delta \varepsilon}},$$

Suponiendo que $L \ll \Omega$, vemos entonces que va a haber un efecto importante si $\Omega - C$ es suficientemente distinto de Ω .

Vemos que la presencia de las moléculas de crowding aumenta la probabilidad de que el ligando esté pegado al receptor (claro que esto no tiene en cuenta cómo explorar el espacio cuando está todo ocupado).

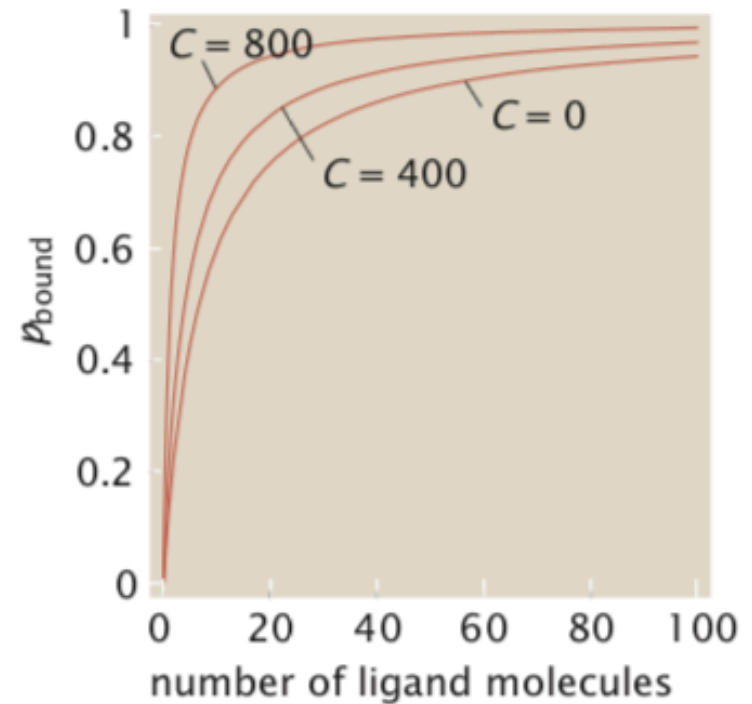


Figure 14.7: Probability of a protein binding site being occupied by a ligand for a number of different concentrations of the crowding molecules. The reaction volume is $\Omega = 1000$ and $\Delta \varepsilon_L = -5 k_B T$.

Recordemos la figura del principio.

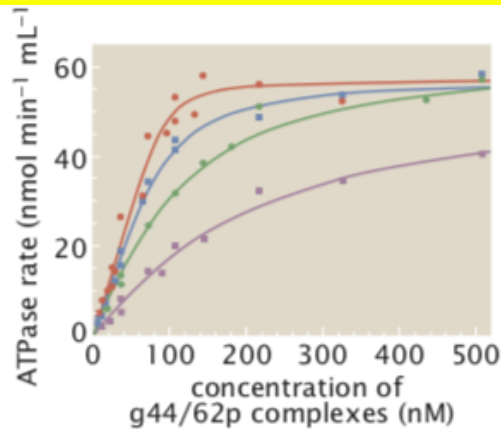


Figure 14.3: ATPase rate associated with T4 DNA replication. The different curves measure the ATPase rate as a function of the g44/62p (clamp loader) protein concentration as measured using different concentrations of polyethylene glycol as a crowding agent. The concentrations of polyethylene glycol going from the bottom to the top curve are 0, 2.5, 5, and 7.5 weight percent. (Adapted from T. C. Jarvis et al., *J. Biol. Chem.* 265:15160, 1990.)

Las curvas teóricas que se muestran en esa figura se ajustan a curvas de unión a partir de las cuales se puede determinar la constante de disociación en función de la concentración de polietilenglicol.

Acá están las constantes de disociación resultantes:

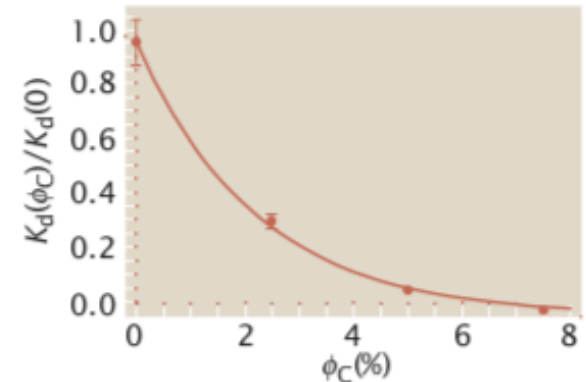


Figure 14.8: Dissociation constant as a function of crowding. Comparison of theory (full line) and experiment (filled circles) for binding in the presence of crowding agents. Measured values of the dissociation constant for g44/62p and g45p components of T4 DNA replication complex as a function of polyethylene glycol concentration measured in percent volume fraction. The theoretical curve was obtained by fitting Equation 14.6 for the effective size ratio r of the protein components to polyethylene glycol 12000 (i.e., with a molecular mass of 12 kDa) molecules. (Adapted from T. C. Jarvis et al., *J. Biol. Chem.* 265:15160, 1990.)

Cualitativamente, lo que muestran estos datos es que a medida que aumenta la concentración de las moléculas de crowding (polietilenglicol), la constante de disociación disminuye, es decir, la reacción de unión es más favorable que la de "desunión".

En este caso el modelo fue un poco más sofisticado. Teniendo en cuenta que las moléculas de polietilenglicol son considerablemente más pequeñas (12 kDa) que el complejo proteico (producto genético 44/62 con una masa de 164 kDa) cuya unión se ve afectada. Para tener en cuenta la diferencia de tamaño, en el modelo se supone entonces que la proteína ocupa r cajas, mientras que una molécula de polietilenglicol ocupa una sola. Para simplificar aún más la combinatoria, se supone que el volumen de reacción se divide en cajas grandes, cada una de las cuales consta de r cajas más pequeñas, como se muestra en la figura:

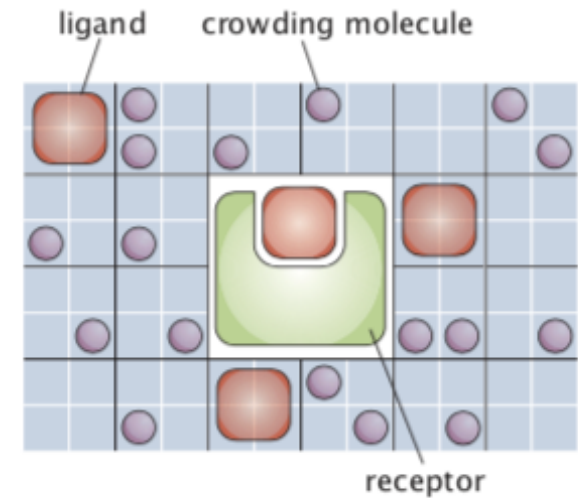


Figure 14.9: Lattice model for large ligands. This lattice model describes binding in the presence of crowding agents where the size of the crowder is different from that of the ligands. This is represented by using different-size boxes for the crowder and the ligand.

Las proteínas ocupan las cajas más grandes, mientras que el polietilenglicol ocupa las cajas más pequeñas desocupadas. Un cálculo más detallado permitiría que la proteína ocupara cualquier región con r cajitas dentro, no solo las marcadas en la figura.

En este caso, la función de partición de la solución “sola” con L ligandos y C moléculas de crowding es:

$$Z_{\text{sol}}(L, C) = \frac{\Omega!}{(\Omega - L)!L!} \frac{(r\Omega - rL)!}{(r\Omega - rL - C)!C!} e^{-\beta L \varepsilon_L^{\text{sol}}} e^{-\beta C \varepsilon_C^{\text{sol}}}$$

Donde Ω es el número de cajas “grandes” (r chicas).

Procediendo como antes, la probabilidad de que el receptor tenga un ligando pegado es:

$$p_{\text{bound}} = \frac{1}{1 + \frac{Z_{\text{sol}}(L, C)}{Z_{\text{sol}}(L - 1, C)} e^{\beta \Delta \varepsilon_L}} = \frac{1}{1 + \frac{\Omega}{L} (1 - \phi_C)^r e^{\beta \Delta \varepsilon_L}}$$

Donde: $\phi_C \equiv C / r\Omega$. La última igualdad, por otro lado, corresponde a suponer $L \ll \Omega$.

También se usó $(N + r)! / N! \approx N^r$ que vale para $r \ll N$.

Dado que: $p_{\text{bound}} = \frac{[L]/K_d}{1 + ([L]/K_d)}$ resulta

$$\frac{K_d(\phi_C)}{K_d(\phi_C = 0)} = (1 - \phi_C)^r.$$

En este caso, la función de partición de la solución “sola” con L ligandos y C moléculas de crowding es:

$$Z_{\text{sol}}(L, C) = \frac{\Omega!}{(\Omega - L)!L!} \frac{(r\Omega - rL)!}{(r\Omega - rL - C)!C!} e^{-\beta L \varepsilon_L^{\text{sol}}} e^{-\beta C \varepsilon_C^{\text{sol}}}$$

Donde Ω es el número de cajas “grandes” (r chicas).

Procediendo como antes, la probabilidad de que el receptor tenga un ligando pegado es:

$$p_{\text{bound}} = \frac{1}{1 + \frac{Z_{\text{sol}}(L, C)}{Z_{\text{sol}}(L - 1, C)} e^{\beta \Delta \varepsilon_L}} = \frac{1}{1 + \frac{\Omega}{L} (1 - \phi_C)^r e^{\beta \Delta \varepsilon_L}}$$

Donde: $\phi_C \equiv C/r\Omega$. La última igualdad, por otro lado, corresponde a suponer $L \ll \Omega$.

También se usó $(N + r)!/N! \approx N^r$ que vale para $r \ll N$.

Dado que: $p_{\text{bound}} = \frac{[L]/K_d}{1 + ([L]/K_d)}$ resulta

$$\frac{K_d(\phi_C)}{K_d(\phi_C = 0)} = (1 - \phi_C)^r.$$

En esta figura:

En esta figura (que ya vimos) se comparó la fórmula:

$$p_{\text{bound}} = \frac{1}{1 + \frac{Z_{\text{sol}}(L, C)}{Z_{\text{sol}}(L-1, C)} e^{\beta \Delta \varepsilon_L}} = \frac{1}{1 + \frac{\Omega}{L} (1 - \phi_C)^r e^{\beta \Delta \varepsilon_L}}$$

con los datos del complejo proteico g44 / 62p del bacteriófago T4. La curva teórica se obtiene determinando el valor, r , que mejor ajusta los datos experimentales. A partir del conocimiento previo de las masas de los ligandos y de los crowders, se podría suponer que r , que es la relación de tamaños entre ligandos y crowders, sería $r \approx 15$. El fiteo da $r = 45$.

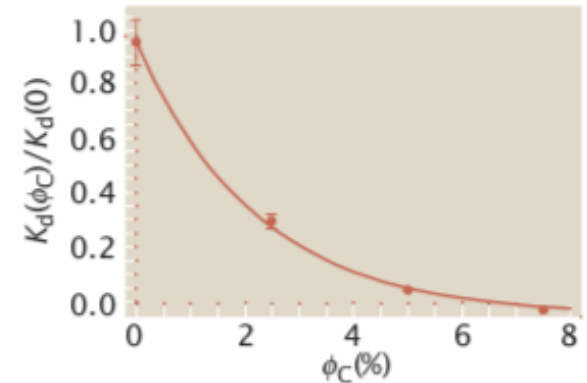


Figure 14.8: Dissociation constant as a function of crowding. Comparison of theory (full line) and experiment (filled circles) for binding in the presence of crowding agents. Measured values of the dissociation constant for g44/62p and g45p components of T4 DNA replication complex as a function of polyethylene glycol concentration measured in percent volume fraction. The theoretical curve was obtained by fitting Equation 14.6 for the effective size ratio r of the protein components to polyethylene glycol 12000 (i.e., with a molecular mass of 12 kDa) molecules. (Adapted from T. C. Jarvis et al., *J. Biol. Chem.* 265:15160, 1990.)

Todos éstos son cálculos de situaciones en equilibrio. La función de partición puede usarse para calcular varias funciones de estado que se introducen en termodinámica, por ejemplo, la energía libre.

Usando la energía libre se puede calcular el trabajo que se produce a partir de una variación y, sabiendo la variación de volumen, determinar la presión, en particular, la presión osmótica, en un caso con “crowding”.

Ejemplo de presión osmótica y crowding

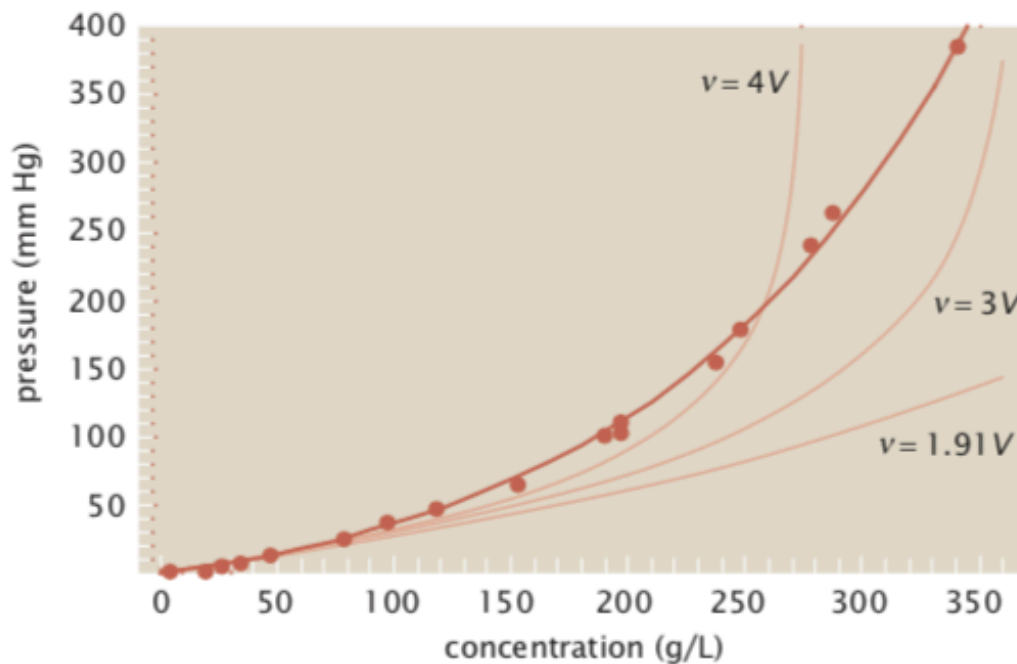


Figure 14.11: Osmotic pressure of a concentrated solution of hemoglobin at 0°C. The filled circles are the experimental data points. The light red lines are predictions of the lattice gas, while the full red line is the pressure of a gas of hard spheres as described by Equation 14.11, with each sphere having a volume V corresponding to a diameter of 5.8 nm. The labels on the lines indicate the volume of a single box in the lattice model given in Equation 14.10. (Data taken from P. D. Ross and A. P. Minton, *J. Mol. Biol.* 112:437, 1977.)

En este caso usamos el modelo de lattice con sitios ocupados por la hemoglobina y sitios libres (ocupados por el solvente).

Si H es el número de moléculas de hemoglobina, Z es:

$$Z_{\text{sol}}(H, \Omega) = \frac{\Omega!}{H!(\Omega - H)!} e^{-\beta H \varepsilon_H^{\text{sol}}},$$

La energía libre y Z se relacionan por:

$$G = -k_B T \ln Z_{\text{sol}},$$

Calculando la variación que se produce en G cuando el volumen total disminuye en una “celda” de volumen v , podemos determinar p a partir de:

$$p v = G(\Omega - 1) - G(\Omega) = -k_B T \ln \frac{Z_{\text{sol}}(H, \Omega - 1)}{Z_{\text{sol}}(H, \Omega)},$$

Por lo que p resulta:

$$p = \frac{k_B T}{v} \ln \frac{\Omega}{\Omega - H},$$

o, equivalentemente,

$$p = -\frac{k_B T}{v} \ln(1 - [H]v). \quad \text{con } [H] = H/(\Omega v)$$

Con esto vimos cómo el aglomeramiento afecta el equilibrio.

Más adelante veremos cómo afecta la dinámica (más específicamente el transporte), algo que ya vimos un poco.