

ENTROPY IN BIOLOGY

Charles S. Peskin

Spring 2020

Lecture 2

Osmotic equilibrium minimizes Gibbs free energy

Stirling's formula

$$\text{Entropy} = k \log (\# \text{ of microstates})$$

Osmotic system with a finite number of molecules

Entropy change when one molecule
crosses the membrane

Markov chain model for osmotic fluctuations
Homework: simulate the fluctuations

Ensemble of finite osmotic systems

Entropy & equilibrium of the ensemble
Ratio of rate constants of the Markov chain

Macroscopic osmotic flow

Osmotic equilibrium at fixed temperature T
and at fixed pressures p^A, p^B minimizes
the Gibbs free energy

$$G = U^A - TS^A + p^A V^A \\ + U^B - TS^B + p^B V^B$$

where, for $I = A, B$

$$U^I = N_0^I u_0(T) + N_1^I u_1(T)$$

$$S^I = N_0^I (\sigma_0(T) - k \log X_0^I) \\ + N_1^I (\sigma_1(T) - k \log X_1^I)$$

$$V^I = N_0^I v_0 + N_1^I v_1$$

In the foregoing, all quantities are constant
 except

$$N_0^A, N_0^B$$

and of course $X_0^A, X_1^A, X_0^B, X_1^B$ depend on
 N_0^A, N_0^B .

/3

Also, N_0^A, N_0^B satisfy the constraint

$$N_0^A + N_0^B = N_0$$

where N_0 is the total # of molecules of solvent in the system. Therefore, we can think of G as a function of N_0^B only. When we differentiate, we have

$$\frac{dN_0^A}{dN_0^B} = -1$$

Note that

$$U^A + U^B = (N_0^A + N_0^B) u_0(T) + (N_1^A + N_1^B) u_1(T)$$

and this is constant.

Also

$$\frac{dS^I}{dN_0^I} = \sigma_0(T) - k \log X_0^I$$

$$-k \left(\frac{N_0^I}{X_0^I} \frac{dX_0^I}{dN_0^I} + \frac{N_1^I}{X_1^I} \frac{dX_1^I}{dN_0^I} \right)$$

$$= \sigma_0(T) - k \log X_0^I$$

~~$$-k (N_0^I + N_1^I) \frac{d}{dN_0^I} (X_0^I + X_1^I)$$~~

and therefore

$$\frac{d}{dN_0^B} (S^A + S^B) = - \frac{dS^A}{dN_0^A} + \frac{dS^B}{dN_0^B}$$

$$= -k \log \frac{X_0^B}{X_0^A}$$

Finally

$$\frac{d}{dN_0^B} (p^A V^A + p^B V^B) = (-p^A + p^B) v_0$$

Then, putting everything together

$$\frac{dG}{dN_0^B} = kT \log \frac{x_0^B}{x_0^A} + (p^B - p^A) v_0$$

Setting this equal to zero gives the condition of osmotic equilibrium found above

To check that this is indeed a minimum,

$$\begin{aligned} \frac{d^2 G}{d(N_0^B)^2} &= \frac{kT}{x_0^B} \frac{\partial x_0^B}{\partial N_0^B} - \frac{kT}{x_0^A} \frac{\partial x_0^A}{\partial N_0^A} \frac{\partial N_0^A}{\partial N_0^B} \\ &= \frac{kT}{x_0^B} \frac{\partial}{\partial N_0^B} \left(\frac{N_0^B}{N_0^B + N_1^B} \right) + \frac{kT}{x_0^A} \frac{\partial}{\partial N_0^A} \left(\frac{N_0^A}{N_0^A + N_1^A} \right) \\ &> 0 \end{aligned}$$

so it is indeed a minimum, and since G is convex, the minimum is unique.

Stirling's formula

The starting point is the formula

$$n! = \int_0^{\infty} t^n e^{-t} dt$$

for which we can give two proofs

Proof # 1

$$\text{Let } f(a, n) = \int_0^{\infty} t^n e^{-at} dt$$

$$= \left(-\frac{\partial}{\partial a}\right)^n \int_0^{\infty} e^{-at} dt$$

$$= \left(-\frac{\partial}{\partial a}\right)^n \frac{1}{a} = n! \left(\frac{1}{a}\right)^{n+1}$$

Now set $a=1$ to obtain

$$\int_0^{\infty} t^n e^{-t} dt = n!$$

as claimed.

/7

proof #2, by induction

To get started, consider $n=0$:

$$\int_0^{\infty} e^{-t} dt = 1 = 0!$$

Now for $n \geq 1$, suppose the claim is valid for $n-1$, so

$$(n-1)! = \int_0^{\infty} t^{n-1} e^{-t} dt$$

$$= \frac{1}{n} \int_0^{\infty} \frac{d(t^n)}{dt} e^{-t} dt$$

$$= \frac{1}{n} \left(\cancel{t^n e^{-t}} \Big|_0^{\infty} + \int_0^{\infty} t^n e^{-t} dt \right)$$

since $n \geq 1$

and we conclude that

$$n! = \int_0^{\infty} t^n e^{-t} dt$$

/∞

Now to "derive" Stirling's formula, note that

$$g_n(t) = t^n e^{-t} = e^{-(t - n \log t)}$$

and the function $t - n \log t$ has its minimum at the solution of

$$1 - \frac{n}{t} = 0 \quad \Rightarrow \quad t = n$$

and the second derivative of $t - n \log t$ is

$$\frac{n}{t^2}$$

which has the value $\frac{1}{n}$ at $t = n$

Therefore

$$g_n(t) \approx e^{-(n - n \log n)} e^{-\frac{1}{2n}(t-n)^2}$$

and from this, extending the limits of integration to $\pm\infty$, we get

$$n! \approx e^{-(n - n \log n)} \int_{-\infty}^{\infty} e^{-\frac{1}{2n}(t-n)^2} dt$$

and

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2n}(t-n)^2} dt = \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx \sqrt{n}$$

$$= \sqrt{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)} dx dy} \sqrt{n}$$

$$= \int_0^{\infty} e^{-\frac{1}{2}r^2} r dr \sqrt{2\pi n} = \sqrt{2\pi n}$$

Thus

$$n! \approx \sqrt{2\pi n} e^{-(n - n \log n)}$$

$$= \sqrt{2\pi n} e^{-n} n^n = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

$$\log n! = \log \left(\frac{n}{e}\right)^n + \frac{1}{2} \log(2\pi n)$$

$$O(n \log n) \quad O(\log n)$$

Now consider the formula

$$S = -k (N_0 \log X_0 + N_1 \log X_1)$$

$$= +k \left(N_0 \log \frac{N_0 + N_1}{N_0} + N_1 \log \frac{N_0 + N_1}{N_1} \right)$$

$$= k \log \left(\left(\frac{N_0 + N_1}{N_0} \right)^{N_0} \left(\frac{N_0 + N_1}{N_1} \right)^{N_1} \right)$$

$$= k \log \left(\frac{(N_0 + N_1)^{N_0 + N_1}}{N_0^{N_0} N_1^{N_1}} \right)$$

$$= k \log \frac{\left(\frac{N_0 + N_1}{e} \right)^{N_0 + N_1}}{\left(\frac{N_0}{e} \right)^{N_0} \left(\frac{N_1}{e} \right)^{N_1}}$$

$$\approx k \log \frac{(N_0 + N_1)!}{(N_0)! (N_1)!}$$

↗

of arrangements of N_0 molecules of one type and N_1 molecules of another type in $N_0 + N_1$ pigeonholes

Let ΔS be the entropy change in the system starting from the state N_0^A, N_0^B and ending at the state $N_0^A - 1, N_0^B + 1$

$$\Delta S = k \log \frac{(N_0^A - 1 + N_1^A)! (N_0^B + 1 + N_1^B)!}{(N_0^A - 1)! (N_1^A)! (N_0^B + 1)! (N_1^B)!}$$

$$- k \log \frac{(N_0^A + N_1^A)! (N_0^B + N_1^B)!}{(N_0^A)! (N_1^A)! (N_0^B)! (N_1^B)!}$$

$$= k \log \frac{N_0^A}{N_0^A + N_1^A} \frac{N_0^B + 1 + N_1^B}{N_0^B + 1}$$

$$= k \log \frac{(X_0^A)_{\text{initial}}}{(X_0^B)_{\text{final}}}$$

$$= -k \log \frac{(X_0^B)_{\text{final}}}{(X_0^A)_{\text{initial}}}$$

$$\Delta G = v_0 (p_B - p_A) + kT \log \frac{(X_0^B)_{\text{final}}}{(X_0^A)_{\text{initial}}}$$

Sign changed because $\Delta G = v_0 (p_B - p_A) - T \Delta S$

Now consider the fluctuations of the osmotic system in which there are fixed numbers of solute molecules

$$N_1^A, N_1^B$$

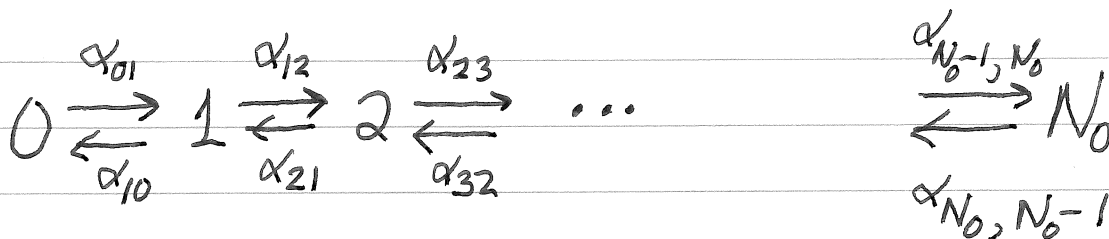
on the two sides of the membrane, and a fixed total number of solvent molecules N_0 in the whole system. Let

$$N_0^A = N_0 - n, \quad N_0^B = n$$

be the numbers of solvent molecules on the two sides of the membrane. Then the possible states of the system are given by

$$n = 0, 1, \dots, N_0$$

and the state evolves according to a continuous-time Markov chain:



in which α_{ij} is the probability per unit time of a transition from i to j when the system is in state i .

For reasons to be discussed later, we assume that

$$\frac{\alpha_{n,n+1}}{\alpha_{n+1,n}} = e^{-\frac{\Delta G_{n,n+1}}{kT}} = e^{-\Theta_{n,n+1}}$$

$$\alpha_{n,n+1} - \alpha_{n+1,n} = -\gamma \frac{\Delta G_{n,n+1}}{kT} = -\gamma \Theta_{n,n+1}$$

where we have set

$$\Theta_{n,n+1} = \frac{\Delta G_{n,n+1}}{kT}$$

and where γ is a constant with units of reciprocal time. We then have the pair of eqns

$$\alpha_{n,n+1} - \alpha_{n+1,n} = -\gamma \Theta_{n,n+1}$$

$$\alpha_{n,n+1} - e^{-\Theta_{n,n+1}} \alpha_{n+1,n} = 0$$

and it follows that

$$\alpha_{n,n+1} = \frac{\gamma \Theta_{n,n+1} e^{-\Theta_{n,n+1}}}{1 - e^{-\Theta_{n,n+1}}} = \frac{\gamma \Theta_{n,n+1}}{e^{\Theta_{n,n+1}} - 1}$$

$$\alpha_{n+1,n} = \frac{\gamma \Theta_{n,n+1}}{1 - e^{-\Theta_{n,n+1}}}$$

Note that $\alpha_{n,n+1}$ and $\alpha_{n+1,n}$ are both positive, regardless of the sign of $\theta_{n,n+1}$

and when $\theta_{n,n+1} = 0$, $\alpha_{n,n+1} = \alpha_{n+1,n} = \gamma$

Also, if we set

$$\theta_{n+1,n} = -\theta_{n,n+1}$$

then

$$\alpha_{n+1,n} = \frac{-\gamma \theta_{n+1,n}}{1 - e^{\theta_{n+1,n}}} = \frac{\gamma \theta_{n+1,n}}{e^{\theta_{n+1,n}} - 1}$$

So $\alpha_{n+1,n}$ depends on $\theta_{n+1,n}$ in exactly

the same way as $\alpha_{n,n+1}$ depends on $\theta_{n,n+1}$

Finally, as a check, we note that

$$\begin{aligned} \frac{\alpha_{n,n+1}}{\alpha_{n+1,n}} &= \frac{1 - e^{-\theta_{n,n+1}}}{e^{\theta_{n,n+1}} - 1} = \frac{(e^{\theta_{n,n+1}} - 1) e^{-\theta_{n,n+1}}}{(e^{\theta_{n,n+1}} - 1)} \\ &= e^{-\theta_{n,n+1}} \quad \text{as required} \end{aligned}$$

and

$$\alpha_{n,n+1} - \alpha_{n+1,n} = \gamma \theta_{n,n+1} \left(\frac{1}{e^{\theta_{n,n+1}} - 1} - \frac{1}{1 - e^{-\theta_{n,n+1}}} \right)$$

$$= \gamma \theta_{n,n+1} \left(\frac{1}{e^{\theta_{n,n+1}} - 1} - \frac{e^{\theta_{n,n+1}}}{e^{\theta_{n,n+1}} - 1} \right)$$

$$= -\gamma \theta_{n,n+1}$$

as required.

Homework

Write a computer program to simulate the Markov chain described above.

Plot some sample paths of $n(t)$.

Evaluate

$$\bar{n} = \frac{1}{T} \int_0^T n(t) dt$$

$$\overline{n^2} = \frac{1}{T} \int_0^T n^2(t) dt$$

over large enough T for the results to be reliable.

Compare \bar{n} to the value that would be expected from macroscopic theory.

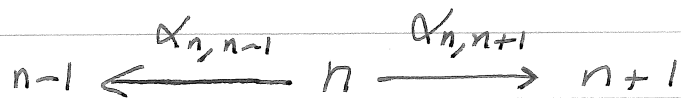
How do the following scale with system size?

$$\sqrt{\overline{n^2} - (\bar{n})^2}, \quad \frac{\sqrt{\overline{n^2} - (\bar{n})^2}}{\bar{n}}$$

Methodology for homework

Use event-driven simulation

In a typical state n , there are two transitions possible:



Do the following:

$$T_+ = -\log(\text{rand}) / \alpha_{n,n+1}$$

$$T_- = -\log(\text{rand}) / \alpha_{n,n-1}$$

if $(T_- < T_+)$

$$n = n - 1$$

$$t = t + T_-$$

else

$$n = n + 1$$

$$t = t + T_+$$

end

In the above sketch of a program:

"rand" in MATLAB gives a random number chosen independently from the

uniform distribution on $(0, 1)$. (Each

call to "rand" gives a different random

number.) Therefore, $-\log(\text{rand})$

gives an exponentially distributed random

number with mean 1.

When computing mean values, remember

that the system is spending variable amounts

of time in each state.

Equilibrium of an ensemble of osmotic microsystems

Each microsystem is characterized by the integer

$$n = \# \text{ of solvent molecules in compartment B}$$

where the possible values of n are

$$n = 0 \dots N_0$$

Let $G(n)$ be the Gibbs free energy in the state n

Let $p(n)$ be the probability of state n at equilibrium, i.e., $p(n)$ is the fraction of members of the ensemble in the state n . We claim that the Gibbs free energy of the whole ensemble divided by the number of members of the ensemble (for a large ensemble) is given by

$$* \quad G = \sum_{n=0}^{N_0} p(n) G(n) + kT \sum_{n=0}^{N_0} p(n) \log p(n)$$

Now we minimize G subject to the constraint

$$\text{that} \quad \sum_{n=0}^{N_0} p(n) = 1 \quad \Rightarrow \quad \sum_{n=0}^{N_0} \delta p(n) = 0$$

We get

$$0 = \delta G = \sum_{n=0}^{N_0} \delta p(n) G(n) + kT \sum_{n=0}^{N_0} \delta p(n) \log p(n)$$

$$+ kT \sum_{n=0}^{N_0} \cancel{p(n)} \frac{\delta p(n)}{\cancel{p(n)}}$$

$$= \sum_{n=0}^{N_0} \delta p(n) (G(n) + kT \log p(n))$$

Since $\delta p(n)$ is arbitrary except that its sum is zero

$$G(n) + kT \log p(n)$$

must be independent of n . It follows that

$$e^{-\frac{G(n') - G(n)}{kT}} = \frac{p(n')}{p(n)} \quad \text{for any pair } n', n$$

In particular

$$e^{-\left(\frac{\Delta G_{n,n+1}}{kT}\right)} = \frac{p(n+1)}{p(n)}$$

This explains why we had to set

$$\frac{\alpha_{n,n+1}}{\alpha_{n+1,n}} = e^{-\frac{\Delta G_{n,n+1}}{kT}}$$

so that we would have

$$\frac{\alpha_{n,n+1}}{\alpha_{n+1,n}} = \frac{p(n+1)}{p(n)}$$

and therefore

$$p(n) \alpha_{n,n+1} = p(n+1) \alpha_{n+1,n}$$

so that the equilibrium p will also be a time-invariant solution of the Markov chain master equation -

It still needs to be shown that G has the form assumed above

Derivation of (*):

We have

$$G(n) = U - kT \log M(n) + P^A V^A(n) + P^B V^B(n)$$

where

$$M(n) = \# \text{ of microstates when } N_0^A = N_0 - n, N_0^B = n$$

The particular form of $M(n)$ will not be important here.

Now consider a large ensemble of osmotic system, and let

$$l(n) = \# \text{ of members of the ensemble with } N_0^A = N_0 - n, N_0^B = n$$

$$L = \sum_{n=0}^{N_0} l(n) = \# \text{ of members of the ensemble.}$$

The number of different ways that the ensemble can be in the state l is

$$\frac{L!}{\prod_{n=0}^{N_0} l(n)!}$$

and for each of them there are $M(n)$ microstates for each of the $l(n)$ members of the ensemble that have $N_0^B = n$.

Therefore the total number of microstates associated with the state l of the ensemble is

$$\frac{L!}{\prod_{n=0}^{N_0} l(n)!} \prod_{n=0}^{N_0} (M(n))^{l(n)}$$

The total entropy of the ensemble in the state l is therefore equal to

$$k \log \frac{L!}{\prod_{n=0}^{N_0} l(n)!} + k \sum_{n=0}^{N_0} l(n) \log (M(n))$$

Since the ensemble is large, we can use Stirling's formula to approximate the first term

$$\log \frac{L!}{\prod_{n=0}^{N_0} l(n)!}$$

$$= \log \frac{L^L}{\prod_{n=0}^{N_0} l(n)^{l(n)}}$$

$$= L \log L - \sum_{n=0}^{N_0} l(n) \log l(n)$$

$$= - \sum_{n=0}^{N_0} l(n) \log \frac{l(n)}{L}$$

So the entropy of the ensemble is

$$-k \sum_{n=0}^{N_0} l(n) \log \frac{l(n)}{L} + k \sum_{n=0}^{N_0} l(n) \log M(n)$$

Note that the second term is the sum of the entropies of the individual ensemble members, so the first term is the additional entropy associated with combining these individuals into an ensemble.

Since the non-entropy terms $n G(n)$ are additive over the ensemble, we have the following expression for the Gibbs free energy of the ensemble

$$\sum_{n=0}^{N_0} l(n) G(n) + kT \underbrace{\sum_{n=0}^{N_0} l(n) \log \frac{l(n)}{L}}_{< 0}$$

note sign

Now divide by L and let $p(n) = \frac{l(n)}{L}$

and we get the average Gibbs free energy per ensemble member as

$$G = \sum_{n=0}^{N_0} p(n) G(n) + kT \sum_{n=0}^{N_0} p(n) \log p(n)$$

which is the same as (*).

Osmotic flow in a non-equilibrium case.

We can use the Markov chain model introduced above to evaluate the osmotic flow in the macroscopic situation that the chambers on either side of the membrane are so large that we can regard the mole fractions of solvent on the two sides as constants. Then

$$\Delta G_{n,n+1} = v_0 (p^B - p^A) + kT \log \frac{x_0^B}{x_0^A}$$

independent of n , and therefore

$$\alpha_{n,n+1} - \alpha_{n+1,n} = -\gamma \frac{\Delta G_{n,n+1}}{kT}$$

$$= \frac{\gamma v_0}{kT} \left(p^A - p^B - \frac{kT}{v_0} \log \frac{x_0^B}{x_0^A} \right)$$

This is the expected value of the net number of molecules per unit time crossing the membrane with the direction $A \rightarrow B$ considered positive.

We can convert it to a flow by multiplying by the volume per molecule v_0 :

$$Q = v_0 \frac{\gamma v_0}{kT} \left(p^A - p^B - \frac{kT}{v_0} \log \frac{x_0^B}{x_0^A} \right)$$

Here Q is the mean volume/time of solvent crossing the membrane.

Note that γ may depend on temperature.

The constant $\frac{\gamma v_0^2}{kT}$ is called the

hydraulic permeability of the membrane

Note that Q depends linearly on P , and

that osmotic pressure produces the same

amount of flow as the corresponding actual

pressure. The linearity is a consequence

of the choice $\alpha_{n,n+1} - \alpha_{n+1,n}$ being ~~the~~ linear in $\Delta G_{n,n+1}$.