BE/APh 161: Physical Biology of the Cell

Justin Bois

Caltech

Winter, 2019

This document was prepared at Caltech with financial support from the Donna and Benjamin M. Rosen Bioengineering Center.

4 Statistical mechanics and ligand-receptor binding

In the last lecture, we explored how to mathematize cartoons, mostly where the underlying physics could be described with mass action kinetics. Today, we will learn how to mathematize cartoons where the physical principles involved rest on **statistical mechanics**. We will have in mind an example, ligand receptor binding, as we do this.

4.1 Motivation: ligand-receptor binding

A cartoon for ligand-receptor binding is shown in Fig. 9. We are interested in computing the probability that a given receptor is bound with a ligand. We will call this p_{bound}. We model the receptor as fixed, sitting in a sea of solvent and ligand. Either one or zero ligands may be bound to the receptor at any given time.

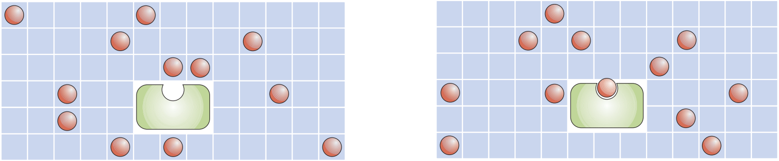

Figure 9: A schematic for ligand-receptor binding. At left, we have a single receptor (in green) and many ligands. Each square in the grid can either be occupied by a ligand or a solvent molecule. In this image, the receptor does not have a ligand bound to it. In the right image, a ligand is bound to the receptor. Figure adapted from Fig. 6.1 of *PBoC2*.

We define a "state," or "microstate," of this system by the configuration of the L ligands among the N available spaces on the grid. Some of these states have the receptor bound. In this case, there are $L - 1$ receptors free to move about the available spaces out in the solvent.

The probability of the receptor being bound is

$$p_{\text{bound}} = \sum_{i \in \text{states with bound receptor}} p_i, \tag{4.1}$$

where p_i is the probability of being in state i. So, in order to compute p_{bound}, we need to compute p_i for a given state. This is where statistical mechanics comes in.

4.2 Derivation of the Boltzmann distribution

We will develop an expression for p_i more generally for any system with an associated set of discrete states. The states are indexed by i, and each has an energy E_i associated with it. We will take an approach along the lines of section 6.1.2 of *PBoC2*, which is a bit unconventional for statistical physics textbooks. We will maximize **informational entropy** in our treatment, following E. T. Jaynes, *Phys. Rev.*, **106**, 620–630, 1957. The abstract of that paper very cleanly and clearly captures the notion of what we are trying to do here.

> Information theory provides a constructive criterion for setting up probability distributions on the basis of partial knowledge, and leads to a type of statistical inference which is called the maximum-entropy estimate. It is the least biased estimate possible on the given information; i.e., it is maximally noncommittal with regard to missing information. If one considers statistical mechanics as a form of statistical inference rather than as a physical theory, it is found that the usual computational rules, starting with the determination of the partition function, are an immediate consequence of the maximum-entropy principle. In the resulting "subjective statistical mechanics," the usual rules are thus justified independently of any physical argument, and in particular independently of experimental verification; whether or not the results agree with experiment, they still represent the best estimates that could have been made on the basis of the information available.
>
> It is concluded that statistical mechanics need not be regarded as a physical theory dependent for its validity on the truth of additional assumptions not contained in the laws of mechanics (such as ergodicity, metric transitivity, equal a priori probabilities, etc.). Furthermore, it is possible to maintain a sharp distinction between its physical and statistical aspects. The former consists only of the correct enumeration of the states of a system and their properties; the latter is a straightforward example of statistical inference.

Indeed, when we perform analysis using statistical mechanics in this class, we will identify the states of the system, assign their energies, and then let the machinations of statistical mechanics do the rest.

4.2.1 The Shannon entropy

The problem of specifying p_i is really open-ended. As Jaynes suggested, we can use maximum-entropy principles to derive an expression for p_i. The entropy he is talking about is the **Shannon entropy**, named after Claude Shannon, who published its mathematical form in 1948, also known as the **informational entropy**. I will state the definition of the entropy associated with a discrete probability distribution, and then give a short discussion on what it means intuitively.

$$S = -K\sum_i p_i \ln p_i, \tag{4.2}$$

where K is an arbitrary positive constant. It is understood that all p_i's are nonnegative and that $p_i \ln p_i \to 0$ as p_i tends toward zero.

We can think of entropy as a measure of ignorance, or of unbiasedness. For example, an unbiased coin will give heads in half of the flips, so the probability of getting heads is $p_h = 1/2$. We can choose K such that

$$S = -\sum_i p_i \log_2 p_i = -p_h \log_2 p_h - (1-p_h)\log_2(1-p_h). \tag{4.3}$$

So, if $p_h = 1/2$, $S = 1$ bit, where a "bit" is the unit of entropy when we have chosen K as we have. Now, let's say $p_h = (1+\varepsilon)/2$, where $\varepsilon \in [-1,1]$. Now, we have

$$\begin{aligned}
S &= -\frac{1+\varepsilon}{2}\log_2\frac{1+\varepsilon}{2} - \frac{1-\varepsilon}{2}\log_2\frac{1-\varepsilon}{2} \\
&= -\frac{1}{2}\log_2\frac{(1+\varepsilon)(1-\varepsilon)}{4} - \frac{\varepsilon}{2}\log_2\frac{1+\varepsilon}{1-\varepsilon}
\end{aligned}$$

$$= 1 - \log_2(1-\varepsilon^2) - \frac{\varepsilon}{2}\,\log_2 \frac{1+\varepsilon}{1-\varepsilon}$$

$$= 1 - \log_2(1-\varepsilon^2) - \frac{|\varepsilon|}{2}\,\log_2 \frac{1+|\varepsilon|}{1-|\varepsilon|}. \tag{4.4}$$

Looking at the three terms, we have a constant plus two monotonically decreasing functions of $|\varepsilon|$. Further, if $|\varepsilon| = 1$, we get $S = 0$. So, the maximal entropy is when ε, the bias of the coin, is zero. The entropy is minimal when $|\varepsilon| = 1$, which means that we know the outcome of the coin toss ahead of time.

Now, imagine that instead of flipping a fair coin (which has two sizes), we roll a fair 8-sided die. The entropy associated with the probability distribution for the die is

$$S = -\sum_i p_i \log_2 p_i = -8\left(\frac{1}{8}\,\log_2 \frac{1}{8}\right) = 3 \text{ bits}. \tag{4.5}$$

So, the entropy for a fair 8-sided die is greater than that of a fair coin. This makes sense; we are more ignorant as to the result we would expect from an 8-sided die than from a two-sided coin.

It turns out that there is only one way to define entropy that satisfies a set of *desiderata*, or desired qualities about entropy, our measure of ignorance, or unbiasedness. These desiderata are, loosely,

1. The entropy is continuous in p_i.
2. If all p_i are equal, the entropy is monotonic in p_i. (Thus, the probability distribution describing the outcomes of a roll of a fair 8-sided die should have greater entropy than that describing a fair coin flip.)
3. Arbitrary grouping of events does not change the entropy (the so-called composition law).

Shannon proved that the *only* function that has these properties is in fact the Shannon entropy, equation 4.2.

4.2.2 The maximal entropy distribution

To be maximally unbiased, or to use *only* the information we have about a system to infer p_i, we must choose p_i that maximizes the entropy. To do this, we differentiate the entropy with respect to p_i and set the derivative equal to zero.

$$\frac{\partial S}{\partial p_j} = -K\,\frac{\partial}{\partial p_j}\sum_i p_i \ln p_i = -K(1+\ln p_j) = 0 \quad \Rightarrow \quad p_j = \mathrm{e}^{-1}. \tag{4.6}$$

I put this equation in gray because this is *not* what we should do! Clearly this cannot be right, since the probability distribution is not normalized, i.e., $\sum_i p_i \neq 1$.

So, we need to do a *constrained* maximization. Specifically, we need to impose the constraint that $\sum_i p_i = 1$, as is always the case. We impose further that p_i has a well-defined expectation value for the energy,

$$\langle E\rangle = \sum_i p_i E_i. \tag{4.7}$$

To impose the constraints in the maximization problem, we use the **method of Lagrange multipliers**, which is described on pages 254–255 in *PBoC2*. The idea is that we add zero to $S(p_i)$, where the “zero” we add is defined by the constraints, with a multiplier and then minimize that function over p_i and the multipliers. We call this function the Lagrangian.

$$\begin{aligned}\mathcal{L}(p_i, \alpha, \beta) &= S + \alpha \left(1 - \sum_i p_i\right) + \beta \left(\langle E\rangle - \sum_i p_i E_i\right) \\ &= -K \sum_i p_i \ln p_i + \alpha \left(1 - \sum_i p_i\right) + \beta \left(\langle E\rangle - \sum_i p_i E_i\right), \qquad (4.8)\end{aligned}$$

where α and β are the Lagrange multipliers. Necessary conditions for $p_i, \alpha, and \beta$ to be maximal are that

$$\frac{\partial \mathcal{L}}{\partial p_j} = 0 \; \forall j, \qquad (4.9)$$

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 0, \qquad (4.10)$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = 0. \qquad (4.11)$$

The last two conditions just mean that the constraints are satisfied, since they reduce to

$$1 - \sum_i p_i = 0, \qquad (4.12)$$

$$\langle E\rangle - \sum_i p_i E_i = 0. \qquad (4.13)$$

Now, if the constraints are affine (meaning that their second derivative with respect to p_i vanishes, which they do) and the entropy is strictly concave (its matrix of second derivatives, called the Hessian, is negative definite), then the necessary conditions are sufficient for optimality. Entry jk of the Hessian is

$$\frac{\partial^2 S}{\partial p_i \partial p_j} = -\frac{\delta_{jk}}{p_j}, \qquad (4.14)$$

where δ_{jk} is the Kronecker delta ($\delta_{jk} = 1$ for $j = k$ and 0 otherwise). This means that the Hessian is diagonal with negative entries, so it is negative definite. Therefore, we need only to solve equations (4.9) through (4.11) to determine the maximum entropy probability distribution, p_i.

Now, we will find where the derivative of the Lagrangian with respect to p_j is zero.

$$\frac{\partial \mathcal{L}}{\partial p_j} = -K(1 + \ln p_j) - \alpha - \beta E_j = 0. \qquad (4.15)$$

Solving for p_j gives

$$p_j = \mathrm{e}^{-\alpha} \, \mathrm{e}^{-\beta E_j}, \qquad (4.16)$$

where we have absorbed constants such that $1 + \alpha/K \to \alpha$ and $\beta/K \to \beta$. Now, using the normalization constraint, we have

$$\sum_i p_i = \mathrm{e}^{-\alpha} \sum_i \mathrm{e}^{-\beta E_i} = 1, \qquad (4.17)$$

so

$$e^{\alpha} = \sum_i e^{-\beta E_i} \equiv Z, \tag{4.18}$$

where we have defined the **partition function** Z. The second constraint, $\langle E \rangle = \sum_i p_i E_i$ is automatically satisfied by definition, so we have arrived at our maximum entropy probability distribution. It is an exponential distribution.

$$p_i = \frac{e^{-\beta E_i}}{Z}. \tag{4.19}$$

4.2.3 Connection to thermodynamics

While we have derived an expression for p_i, we still do not know the physical meaning of the Lagrange multiplier β. We know only that it must have dimensions of inverse energy, since βE_i must be dimensionless. To connect β to physical quantities, we turn to thermodynamics. Thermodynamics deals with observed quantities in large systems. The internal energy is $\langle E \rangle$. We can write the combined first and second law of thermodynamics as

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}\langle E \rangle. \tag{4.20}$$

Then,

$$\frac{\partial S}{\partial \langle E \rangle} = \frac{1}{T}. \tag{4.21}$$

To compute the derivative of the entropy, we first write it in a more convenient form using our derived expression for p_i.

$$\begin{aligned} S &= -K\sum_i p_i \ln p_i = -K\sum_i p_i(-\beta E_i - \ln Z) \\ &= K\beta \sum_i p_i E_i + K \ln Z \sum_i p_i = K\beta\langle E \rangle + K \ln Z. \end{aligned} \tag{4.22}$$

Thus, we have

$$\frac{\partial S}{\partial \langle E \rangle} = K\beta = \frac{1}{T}. \tag{4.23}$$

So, for the Shannon entropy to be equal to the thermodynamic entropy, $K\beta = 1/T$. Thus, $\beta = 1/KT$. When we have equivalence to the thermodynamic entropy, we call the constant K the **Boltzmann constant**, and denote it as k_B or k. The Boltzmann constant has a value of

$$k_B = 1.38 \times 10^{-23}\,\text{J/K} = 4.1\text{ pN-nm}. \tag{4.24}$$

We will also use $\beta \equiv 1/k_B T$ in our calculations, since it turns out to be notationally convenient. Thus, we have

$$p_i = \frac{e^{-E_i/k_B T}}{\sum_i e^{-E_i/k_B T}}. \tag{4.25}$$

The quantity e^{-E_i/k_BT} represents an unnormalized probability and is referred to as a **Boltzmann weight**. Recall that the sum of Boltzmann weights is

$$Z = \sum_i \mathrm{e}^{-E_i/k_BT}, \tag{4.26}$$

which serves as the normalization constant of the probability, is called a **partition function**.

4.3 Back to ligand-receptor binding

We can now return to our ligand-receptor binding problem. We know the probability of each state, p_i, and we just need to assign energies to them to compute p_{bound}. Let the energy of a single unbound ligand be E_u and the energy of a bound ligand be E_b. Then the total energy of any state where the receptor is unbound is LE_u, where, as a reminder, L is the total number of ligands. The total energy of any state where the receptor is bound is $E_b + (L-1)E_u$. Then the total statistical weight of all unbound states is equal to the number of states with unbound receptor times the Boltzmann weight of an unbound state, $\mathrm{e}^{-\beta E_u}$.

We can compute the number of states with unbound ligand. The number of ways select L our of N lattice sites to be occupied by ligand is given by the binomial coefficient, $N!/(N-L)!L!$. This is the **multiplicity** of the bound state; i.e., the number of states with the same energy.

It helps to organize everything into a **states and weights table**.

state	**energy**	**multiplicity**	**statistical weight**
receptor unbound	LE_u	$\frac{N!}{(N-L)!L!}$	$\frac{N!}{(N-L)!L!}\,\mathrm{e}^{-\beta LE_u}$
receptor bound	$E_b + (L-1)E_u$	$\frac{N!}{(N-L+1)!(L-1)!}$	$\frac{N!}{(N-L+1)!(L-1)!}\,\mathrm{e}^{-\beta(E_b+(L-1)E_u)}$

For ease of notation, we will denote the appropriate binomial coefficients as Ω_u and Ω_b. Then, we can compute p_{bound} as

$$p_{\text{bound}} = \frac{\Omega_b \mathrm{e}^{-\beta(E_b+(L-1)E_u)}}{\Omega_b \mathrm{e}^{-\beta(E_b+(L-1)E_u)} + \Omega_u \mathrm{e}^{-\beta LE_u}} = \frac{\frac{\Omega_b}{\Omega_u}\,\mathrm{e}^{-\beta(E_b-E_u)}}{1+\frac{\Omega_b}{\Omega_u}\,\mathrm{e}^{-\beta(E_b-E_u)}}. \tag{4.27}$$

Now,

$$\frac{\Omega_b}{\Omega_u} = \frac{N!}{(N-L+1)!(L-1)!}\,\frac{(N-L)!L!}{N!} = \frac{L}{N-L+1}. \tag{4.28}$$

Because $L \ll N$ in a dilute solution,

$$N - L + 1 \approx N \approx N + L. \tag{4.29}$$

Using these approximate expressions allow us to write Ω_b/Ω_u as the **mole fraction** of ligand, x_L.

$$\frac{\Omega_b}{\Omega_u} \approx \frac{L}{N} \approx \frac{L}{N+L} = x_L. \tag{4.30}$$

So,

$$p_i = \frac{x_L \mathrm{e}^{-\beta(E_b-E_u)}}{1 + x_L \mathrm{e}^{-\beta(E_b-E_u)}}. \tag{4.31}$$

We can convert mole fraction to concentration be multiplying by the density of solvent, water in most physiological cases;

$$c_L = \rho_{\mathrm{H_2O}} x_L. \tag{4.32}$$

Now, if we multiply top and bottom of our expression for p_i by unity, represented as $\rho_{\mathrm{H_2O}}/\rho_{\mathrm{H_2O}}$, we get

$$p_i = \frac{c_L \mathrm{e}^{-\beta(E_b - E_u)}/\rho_{\mathrm{H2O}}}{1 + c_L \mathrm{e}^{-\beta(E_b - E_u)}/\rho_{\mathrm{H2O}}}. \tag{4.33}$$

Finally, we define the **dissociation** constant $K_\mathrm{d} = \rho_{\mathrm{H_2o}} \mathrm{e}^{-\beta(E_u - E_b)}$. We arrive at

$$p_\mathrm{bound} = \frac{c_L/K_\mathrm{d}}{1 + c_L/K_\mathrm{d}}. \tag{4.34}$$

This is a common result (called a Langmuir isotherm) that could be seen from what you remember from general chemistry. We can take the probability of a receptor being bound as

$$p_\mathrm{bound} = \frac{c_{LR}}{c_{LR} + c_R}. \tag{4.35}$$

We use the definition of the dissociation constant,

$$K_\mathrm{d} = \frac{c_L c_R}{c_{LR}}, \tag{4.36}$$

to get

$$p_\mathrm{bound} = \frac{c_L c_R/K_\mathrm{d}}{c_L c_R/K_\mathrm{d} + c_R} = \frac{c_L/K_\mathrm{d}}{1 + c_L/K_\mathrm{d}}. \tag{4.37}$$

In deriving this result, we have a clear picture about the physical origin of the dissociation constant. We also have a framework to study cases where we have more complicated states and weights.

4.4 Maximum entropy distributions for other ensembles

We will now use the method of maximum entropy to derive probability distributions when we know other facts about the states.

4.4.1 Given energy and number of particles

Now let's say that we have a system that consists of particles. Each state of the system has a well defined energy, E_i and number of particles, N_i. We should therefore have an expectation value for N_i. Now, we have three constraints for p_i.

$$\sum_i p_i = 1, \tag{4.38}$$

$$\langle E \rangle = \sum_i p_i E_i, \tag{4.39}$$

$$\langle N\rangle = \sum_i p_i N_i. \tag{4.40}$$

We construct our Lagrangian as before, but with a third Lagrange multiplier.

$$\mathcal{L} = -k_B \sum_i p_i \ln p_i + \alpha \left(1 - \sum_i p_i\right) + \beta \left(\langle E\rangle - \sum_i p_i E_i\right) + \gamma \left(\langle N\rangle - \sum_i p_i N_i\right). \tag{4.41}$$

We take the same approach as before.

$$\frac{\partial \mathcal{L}}{\partial p_j} = -k_B(1 + \ln p_j) - \alpha - \beta E_j - \gamma N_j = 0. \tag{4.42}$$

Solving gives

$$p_j = \mathrm{e}^{-\alpha}\, \mathrm{e}^{-\beta E_j - \gamma N_j}, \tag{4.43}$$

where we have again absorbed constants: $-1 - \alpha/k_B \to \alpha$, $\beta/k_B \to \beta$, and $\gamma/k_B \to \gamma$. We use the normalization condition that $\sum_i p_i = 1$ to get

$$\mathrm{e}^{\alpha} = \sum_i \mathrm{e}^{-\beta E_j - \gamma N_j} \equiv Z. \tag{4.44}$$

To find the values of the other Lagrange multipliers that connect the entropy to the thermodynamic entropy, we do the same procedure. We first write the combined first and second law of thermodynamics.

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}\langle E\rangle - \frac{\mu}{T}\,\mathrm{d}\langle N\rangle, \tag{4.45}$$

where μ is the chemical potential of the particles. Thus,

$$\left(\frac{\partial S}{\partial \langle N\rangle}\right)_{\langle E\rangle} = -\frac{\mu}{T}. \tag{4.46}$$

Going back to the expression we wrote for the probability p_i and the partition function,

$$S = -k_B \sum_i p_i \ln p_i = -k_B \sum_i p_i \left(-\beta E_i - \gamma N_i - \ln Z\right) = k_B \ln Z + k_B \beta \langle E_i\rangle + k_B \gamma \langle N\rangle. \tag{4.47}$$

So,

$$\left(\frac{\partial S}{\partial \langle N\rangle}\right)_{\langle E\rangle} = k_B \gamma = -\frac{\mu}{T}. \tag{4.48}$$

Thus, we have $\gamma = -\mu/k_B T$. We again get $\beta = 1/k_B T$ in a similar manner. Thus, we have

$$p_i = \frac{\mathrm{e}^{-\beta(E_i - \mu N_i)}}{\sum_i \mathrm{e}^{-\beta(E_i - \mu N_i)}}. \tag{4.49}$$

4.4.2 A general thermodynamic conjugate pair

We see a pattern here. Let's say that a given state has associated with it an energy E_i, and another arbitrary extensive property X_i with a well-defined expectation value $\langle X\rangle$. Then, the maximum entropy distribution is

$$p_i = \frac{\mathrm{e}^{-\beta E_i - \lambda X_i}}{\mathcal{Z}}, \tag{4.50}$$

where

$$\mathcal{Z} = \sum_i \mathrm{e}^{-\beta E_i - \lambda X_i}, \tag{4.51}$$

with λ being a Lagrange multiplier. To link λ to a physical quantity, it always ends up being the thermodynamic conjugate variable to $\langle X\rangle$ divided by kT. We can have many such extensive properties. So, if we index these properties by k, we have, generally,

$$p_i = \frac{1}{\mathcal{Z}} \exp\left\{-\frac{1}{k_BT}\left(E_i + \sum_k y_k X_k\right)\right\}, \tag{4.52}$$

where y_k denotes the thermodynamic conjugate variable to X_k and

$$\mathcal{Z} = \sum_i \exp\left\{-\frac{1}{k_BT}\left(E_i + \sum_k y_k X_k\right)\right\}. \tag{4.53}$$

4.5 Another look at ligand-receptor binding

Let's take another look at ligand-receptor binding using our new tools. We'll reframe how we look at the system. We focus on the receptor, knowing there is a pool of ligands immediately around it. In the *immediate* vicinity of the receptor, there can only be zero or one ligand. In the latter case, the ligand is bound. So, the number of ligands in the system can fluctuate, so we can define each state to have an energy E_i and a number of ligands, L_i. If μ is the chemical potential of a ligand, then we have a new states and weights table.

state	**energy**	**multiplicity**	**statistical weight**
receptor unbound	E_u	1	$\mathrm{e}^{-\beta E_u}$
receptor bound	E_b	1	$\mathrm{e}^{-\beta(E_b-\mu)}$

We can then readily compute p_bound.

$$p_\text{bound} = \frac{\mathrm{e}^{-\beta(E_b-\mu)}}{\mathrm{e}^{-\beta(E_b-\mu)} + \mathrm{e}^{-\beta E_u}} = \frac{\mathrm{e}^{-\beta(E_b-E_u-\mu)}}{1 + \mathrm{e}^{-\beta(E_b-E_u-\mu)}}. \tag{4.54}$$

Now, as derived in section 6.2.2 of *PBoC2* (we will not derive it here), for a dilute solution, the chemical potential of solute species k is

$$\mu_k = \mu_k^0 + k_BT\ln x_k. \tag{4.55}$$

The chemical potential for the solvent is

$$\mu_{\text{solv}} = \mu^0_{\text{solv}} - k_B T \sum_k x_k. \tag{4.56}$$

If we insert the chemical potential for solute into our expression for p_{bound}, we get

$$p_{\text{bound}} = \frac{x_L \mathrm{e}^{-\beta(E_b - E_u - \mu^0)}}{1 + x_L \mathrm{e}^{-\beta(E_b - E_u - \mu^0)}} = \frac{c_L/K_{\text{d}}}{1 + c_L/K_{\text{d}}} = \frac{c_L}{K_{\text{d}} + c_L}, \tag{4.57}$$

the same expression as before with

$$K_d = \mathrm{e}^{-\beta(E_u + \mu^0 - E_b)}. \tag{4.58}$$

Note that there is a difference in the definition of K_{d}, which is due to the subtle difference in the definition of the energies of the states. In our previous treatment, we defined E_u to be the energy of a ligand when unbound. We tacitly assumed that the energy of the receptor when unbound was zero. Here, E_u is the energy of the receptor when unbound and μ^0 is the energy of a single ligand alone in solution.