Gooey tutorial 3: Aggregation

ICFP M2 – Advanced Biophysics

1 Length distribution of actin filaments

Here we consider the growth and shrinkage of actin filaments, a type of protein fiber used by the cell to regulate its mechanics. We use it as an example of the breaking of the detailed balance condition on the system's statistics.

In our simplified model, we concentrate on a single filament composed of L monomers. The filament is surrounded by monomers in solution with a fixed chemical potential μ, and individual monomers may attach to or detach from the filament, implying changes in the length of the type $L \to L \pm 1$ (Fig. 1). When the protein cofilin is present in the solution, the filament may additionally break at a randomly chosen location along its length. The subsequent step is conditioned by the asymmetry of the filament: its ends are distinct, and known as the "barbed" and "pointed" end respectively. The monomer located at the barbed end is in a different state than all the others due to an internal nonequilibrium dynamics involving the hydrolysis of the cellular fuel ATP. This special state largely protects the barbed end from disassembly. When cofilin breaks the filament, a new barbed end is created that does not benefit from this protection. The filament portion associated with this new barbed end thus disassembles quickly and turns back to monomers.

1.1 Assuming that the energy of a bond between two monomers is ϵ and that the chemical potential of a monomer in solution is μ, show that the equilibrium length distribution for the filament is

$$P^{\mathrm{eq}}(L) = \left[e^{\beta(\epsilon-\mu)} - 1\right] e^{L\beta(\mu-\epsilon)} \qquad \text{for } L \in \mathbb{N}^* \tag{1}$$

1.2 Now consider kinetics of filament elongation: starting from a filament of length L, monomers are stochastically added at a rate 1, and removed at a rate q. Using Eq. (1), compute the value of q that satisfies the detailed balance condition

$$\forall (L, L') \quad P^{\mathrm{eq}}(L) k_{L\to L'} = P^{\mathrm{eq}}(L') k_{L'\to L}, \tag{2}$$

where $k_{L\to L'}$ is the rate at which the system transitions from length L to length L'.

1.3 Write a computer program based on the Gillespie algorithm that simulates the time evolution of the filament.

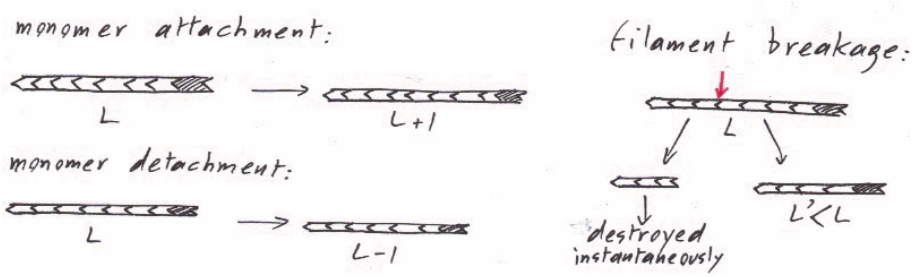


Figure 1: Illustration of the actin filament dynamics described in the text. *Left:* monomer attachment or detachment. *Right:* filament breakage at the location of the red arrow induced by cofilin. The hatched monomers are associated with ATP (in real actin several monomers may simultaneously be associated with ATP, but they do tend to be located in the vicinity of the barbed end).

1.4 Plot the time evolution of the length to make sure everything is going OK, then plot the histogram of the lengths over the course of the simulation. Be careful to assign a statistical weight to each length that is proportional to the time spent with that length.

1.5 Compare the resulting length distribution with Eq. (1). Find a regime where you can collect good statistics within a reasonable computation time. You may assess the quality of your statistics by running the simulation several times and plotting the histogram representing the length distribution. The histogram must be reproducible from one run to the next, and hopefully consistent with the theoretical prediction.

1.6 Now add a new transition to the program to represent the action of cofilin: each bond within the filament now breaks with a rate k. Following this event, keep the part of the filament associated with the original barbed end and destroy the other.

1.7 Argue that the resulting distribution cannot be described by an equilibrium distribution of the type of Eq. (1), even by introducing an effective bond energy.

1.8 Keeping the severing transition, introduce a new transition in the dynamics of the filament that restores the detailed balance condition Eq. (2). This transition may be viewed as an idealization of the coalescence of two filaments in the absence of the ATP-induced filament cap.

1.9 Show numerically that the introduction of this transition restores the equilibrium statistics.

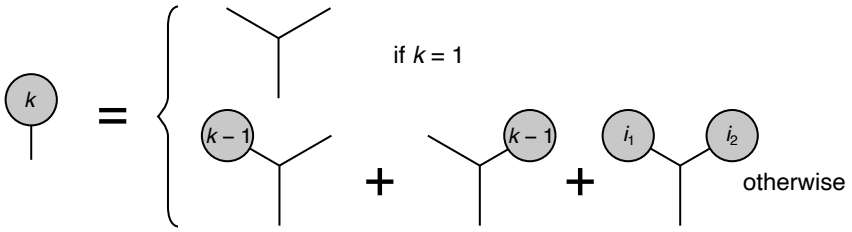
2 Gelation transition: number of aggregates of size k

Here we derive the coefficient N_k introduced in the main lecture as the number of tree-like aggregates made of k particles each carrying z reaction sites. To this effect we assume that we hold one dangling (unreacted) bond in our hand, and compute the number w_k of aggregates of size k that can be built starting from this bond. As we will see below $w_k \neq N_k$ because of a multiple-counting subtlety.

2.1 First establish the recursion equation

$$w_k = z\delta_{k,1} + z(z-1)w_{k-1} + z\frac{(z-1)(z-2)}{2}\sum_{i_1+i_2=k-1} w_{i_1}w_{i_2} + \ldots + z\sum_{i_1+i_2+\ldots+i_z=k-1} w_{i_1}w_{i_2}\ldots w_{i_z}, \tag{3}$$

where the trick is to notice that there are only few ways to make a k-sized aggregate out of smaller aggregates, as illustrated below for $z = 3$:



where dangling black lines picture unreacted bonds and grey circles stand for unspecified aggregates of the size specified in the circle. The different bonds on a monomer are regarded as distinguishable.

2.2 Introducing the generating function $g(x) = \sum_{k=1}^{+\infty} x^k w_k$, use the recursion relation to prove that

$$g(x) = zx\,[1 + g(x)]^{z-1}\,. \tag{4}$$

2.3 Solve this equation in the case $z = 3$ and prove that

$$g(x) = \sum_{k=1}^{+\infty} \frac{(2k)!}{k!(k+1)!}(3x)^k. \tag{5}$$

2.4 Noting that our procedure of growing an aggregate from one of its dangling bonds counts k-sized aggregates as many times as they have dangling bonds, argue that

$$N_k = \frac{(2k)!}{k!(k+2)!}3^k. \tag{6}$$

3 Gelation transition: equilibrium fragmentation rate

From this point on we go back to working with a system with arbitrary z.

3.1 Denoting the energy of a reacted bond by ϵ and using the results from the main lecture, show that the equilibrium concentration of aggregates of size k is

$$c_k^{\text{eq}} = N_k \left(c_1 e^{\beta\epsilon}\right)^k e^{-\beta\epsilon} \tag{7}$$

3.2 Assuming the dynamics of the system is governed by the Smoluchowski equation with rates defined as in the main lecture, show that a system with an equilibrium dynamics must satisfy

$$F_{ij} = K_{ij}\frac{N_i N_j}{N_{i+j}}e^{-\beta\epsilon}. \tag{8}$$

3.3 From this point on now assume a system where all functional groups have the same affinity regardless of their size, imposing

$$K_{ij} = \sigma_i \sigma_j, \tag{9}$$

where $\sigma_i = 2 + i(z-2)$ is the number of dangling bonds of a size-i aggregate, as discussed in the main lecture. Deduce from Eq. (8) that the breaking rate of a single bond is

$$F_{11} = 2e^{-\beta\epsilon} \tag{10}$$

3.4 Use a physical reasoning (rather than a difficult calculation) to argue that for such a system

$$\sum_{i+j=k} F_{ij} = 2(k-1)e^{-\beta\epsilon} \tag{11}$$

3.5 Show that this condition is equivalent to

$$\sum_{i+j=k} (\sigma_i N_i)(\sigma_j N_j) = 2(k-1)N_k \tag{12}$$

and propose a combinatorial interpretation for this equation

4 Gelation transition: extent of reaction equation

We introduce the Ansatz

$$c_k(t) = N_k \left(\frac{\alpha}{z}\right)^{k-1}(1-\alpha)^{\sigma_k}, \tag{13}$$

where $\alpha(t)$ is the extent of reaction, *i.e.*, the probability that a randomly chosen bond is reacted at time t. Note that we chose units of volume such that $\sum_{k=1}^{+\infty} k c_k = 1$.

4.1 By recognizing the number of unreacted bonds per unit volume, use a physical argument to show that

$$\sum_{k=1}^{+\infty} \sigma_k c_k = z(1-\alpha) \tag{14}$$

4.2 If the aggregate concentration satisfies our Ansatz, express the product

$$\frac{N_{i+j}}{N_i N_j} \frac{c_i c_j}{c_{i+j}} \tag{15}$$

as a function of z and α.

4.3 Insert the Ansatz into Smoluchowski's equation and show that it is a solution of the problem if only monomers are present at $t = 0$ and

$$\dot{\alpha} = z(1-\alpha)^2 - e^{-\beta\epsilon}\alpha. \tag{16}$$

Interpret this equation in terms of the reactivity of the bonds.

[Optional] Full solution of the aggregation-fragmentation problem

Solve Eq. (16) to deduce a fully analytical expression for the aggregate concentrations over time. Note that this solution is only valid for the case of bonds with equal reactivity starting from a pure solution of monomers. For other cases, a numerical approach is usually required, for instance using the direct simulation Monte-Carlo method.