Exercises

8.10 **Stochastic cells.**[1] [2] (Biology, Computation) ④

Living cells are amazingly complex mixtures of a variety of complex molecules (RNA, DNA, proteins, lipids, ...) that are constantly undergoing reactions with one another. This complex of reactions has been compared to computation; the cell gets input from external and internal sensors, and through an intricate series of reactions produces an appropriate response. Thus, for example, receptor cells in the retina 'listen' for light and respond by triggering a nerve impulse.

The kinetics of chemical reactions are usually described using differential equations for the concentrations of the various chemicals, and rarely are statistical fluctuations considered important. In a cell, the numbers of molecules of a given type can be rather small; indeed, there is (often) only one copy of the relevant part of DNA for a given reaction. It is an important question whether and when we may describe the dynamics inside the cell using continuous concentration variables, even though the actual numbers of molecules are always integers.

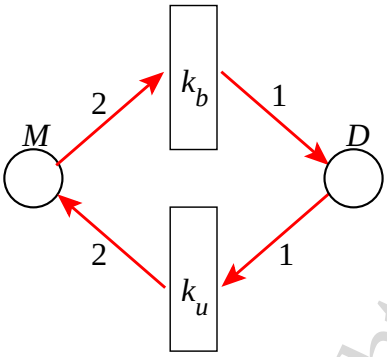

Fig. 8.11 Dimerization reaction. A Petri net diagram for a dimerization reaction, with dimerization rate k_b and dimer dissociation rate k_u.

Consider a dimerization reaction; a molecule M (called the 'monomer') joins up with another monomer and becomes a dimer D: $2M \longleftrightarrow D$. Proteins in cells often form dimers; sometimes (as here) both proteins are the same (homodimers) and sometimes they are different proteins (heterodimers). Suppose the forward reaction rate is k_d and the backward reaction rate is k_u. Figure 8.11 shows this as a Petri net [50] with each reaction shown as a box, with incoming arrows showing species that are consumed by the reaction, and outgoing arrows showing species that are produced by the reaction; the number consumed or produced (the *stoichiometry*) is given by a label on each arrow. There are thus two reactions: the backward unbinding reaction rate per unit volume is $k_u[D]$ (each dimer disassociates with rate k_u), and the forward binding reaction rate per unit volume is $k_b[M]^2$ (since each monomer must wait for a collision with another monomer before binding, the rate is proportional to the monomer concentration squared).

The brackets [.] denote concentrations. We assume that the volume per cell is such that one molecule per cell is 1 nM (10^{-9} moles per liter). For convenience, we shall pick nanomoles as our unit of concentration, so $[M]$ is also the number of monomers in the cell. Assume $k_b = 1\,\mathrm{nM}^{-1}\mathrm{s}^{-1}$ and $k_u = 2\,\mathrm{s}^{-1}$, and that at $t = 0$ all N monomers are unbound.

(a) Continuum dimerization. *Write the differential equation for* $\mathrm{d}M/\mathrm{d}t$ *treating* M *and* D *as continuous variables.* (Hint: Remember that two M molecules are consumed in each reaction.) *What are the equilibrium concentrations for* $[M]$ *and* $[D]$ *for* $N = 2$ *molecules in the cell, assuming these continuous equations and the values above for* k_b *and* k_u*? For* $N = 90$ *and* $N = 10\,100$ *molecules? Numerically solve your differential equation for* $M(t)$

[1] From *Statistical Mechanics: Entropy, Order Parameters, and Complexity* by James P. Sethna, copyright Oxford University Press, 2007, page 178. A pdf of the text is available at pages.physics.cornell.edu/sethna/StatMech/ (select the picture of the text). Hyperlinks from this exercise into the text will work if the latter PDF is downloaded into the same directory/folder as this PDF.

[2] This exercise and the associated software were developed in collaboration with Christopher Myers.

for $N = 2$ and $N = 90$, and verify that your solution settles down to the equilibrium values you found.

For large numbers of molecules in the cell, we expect that the continuum equations may work well, but for just a few molecules there surely will be relatively large fluctuations. These fluctuations are called *shot noise*, named in early studies of electrical noise at low currents due to individual electrons in a resistor. We can implement a Monte Carlo algorithm to simulate this shot noise.[3] Suppose the reactions have rates Γ_i, with total rate $\Gamma_{\rm tot} = \sum_i \Gamma_i$. The idea is that the expected time to the next reaction is $1/\Gamma_{\rm tot}$, and the probability that the next reaction will be j is $\Gamma_j/\Gamma_{\rm tot}$. To simulate until a final time t_f, the algorithm runs as follows.

(1) Calculate a list of the rates of all reactions in the system.

(2) Find the total rate $\Gamma_{\rm tot}$.

(3) Pick a random time $t_{\rm wait}$ with probability distribution $\rho(t) = \Gamma_{\rm tot} \exp(-\Gamma_{\rm tot}\, t)$.

(4) If the current time t plus $t_{\rm wait}$ is bigger than t_f, no further reactions will take place; return.

(5) Otherwise,
   - increment t by $t_{\rm wait}$,
   - pick a random number r uniformly distributed in the range $[0, \Gamma_{\rm tot})$,
   - pick the reaction j for which $\sum_{i<j} \Gamma_i \le r < \sum_{i<j+1} \Gamma_i$ (that is, r lands in the jth interval of the sum forming $\Gamma_{\rm tot}$),
   - execute that reaction, by incrementing each chemical involved by its stoichiometry.

(6) Repeat.

There is one important additional change:[4] the binding reaction rate for M total monomers binding is no longer $k_b M^2$ for discrete molecules; it is $k_b M(M-1)$.[5]

(b) Stochastic dimerization. *Implement this algorithm for the dimerization reaction of part (a). Simulate for $N = 2$, $N = 90$, and $N = 10\,100$ and compare a few stochastic realizations with the continuum solution. How large a value of N do you need for the individual reactions to be well described by the continuum equations (say, fluctuations less than $\pm 20\%$ at late times)?*

Measuring the concentrations in a single cell is often a challenge. Experiments often average over many cells. Such experiments will measure a smooth time evolution even though the individual cells are noisy. Let us investigate whether this ensemble average is well described by the continuum equations.

(c) Average stochastic dimerization. *Find the average of many realizations of your stochastic dimerization in part (b), for $N = 2$ and $N = 90$, and compare with your deterministic solution. How much is the long-term average shifted by the stochastic noise? How large a value of N do you need for the ensemble average of $M(t)$ to be well described by the continuum equations (say, shifted by less than 5% at late times)?*

[3] In the context of chemical simulations, this algorithm is named after Gillespie [45]; the same basic approach was used just a bit earlier in the Ising model by Bortz, Kalos, and Lebowitz [19], and is called *continuous-time Monte Carlo* in that context.

[4] Without this change, if you start with an odd number of cells your concentrations can go negative!

[5] Again $[M] = M$, because we assume one molecule per cell gives a concentration of 1 nM.