

2.3 Molecule counting more generally

Please be warned that this section is much rougher than some of the previous sections.

Section last updated
April 14, 2009.

Chemical signaling is ubiquitous in biological systems (cf Fig 2.8), so it is important to understand its limits. To provide some further motivation, let's think about the regulation of gene expression. We recall that every cell in our bodies has the same DNA. What makes a liver cell different from a neuron in your brain is that it reads out or "expresses" different genes, making different proteins. Importantly, this is not just a discrete choice made once in your lifetime. Given that a set of proteins are being made, the exact amounts of these molecules are constantly being adjusted to match the needs of the cell. This happens also in bacteria, which adjust, for example, the concentrations of the enzymes needed to metabolize different nutrients that might or might not be present in the environment. Indeed, much of what we know about the regulation of gene expression has its roots in work on this sort of metabolic control in bacteria.

should check where this is
discussed first .. maybe with
different rhodopsins?

There are many ways in which gene expression is controlled. As a simple example, note that if we want to regulate the number of proteins in the cell we can change either the rate at which they are made or the rate at which they are degraded, and both these things happen. The synthesis of a protein involves two very different steps, transcription from DNA to messenger RNA and translation from mRNA to protein, and again there is regulation of both processes. All this being said, we will focus our attention on the regulation of transcription, that is the reading of the DNA template to make mRNA.

need some more sketches
here

In order to make mRNA, a complex of proteins (including the RNA polymerase) must bind to the DNA and 'walk' along it, spewing out the mRNA polymer as it walks. In order for all of this to happen, the RNA polymerase has to find the right starting point. One can imagine that this can be inhibited simply by having other proteins bind to nearby sites along the DNA. Alternatively, binding of proteins to slightly different positions near the starting point could help the RNA polymerase to find its way. Both of these things happen: proteins called transcription factors can act both as repressors and as activators of mRNA synthesis. The key step in this regulation is thought to be the binding of the transcription factors to specific sites near the RNA polymerase start site, and these binding sites are specific because the transcription factor protein is selective for particular DNA sequences. Much can be said about the nature of this specificity. For now the important point is that such regulatory systems are, in effect, sensors of the transcription factor concentration.

The binding sites along DNA for the transcription factors have linear dimensions measured in nanometers, perhaps $a \sim 3$ nm. The diffusion constants of proteins in the interior of cells is in the range of $D \sim 1 \mu^2/\text{s}$. Many transcription factors act at nanoMolar concentrations, and it is useful to note that $1 \text{ nM} = 0.6 \text{ molecules}/\mu\text{m}^3$. Putting these together we have $Dac \sim 1.8 \times 10^{-3} \text{ s}^{-1}$. Thus, the Berg–Purcell limit predicts that the smallest changes in transcription factor that can be reliably detected will be in the range

$$\frac{\delta c}{c} \sim \frac{1}{\sqrt{Dac\tau_{\text{avg}}}} \sim \sqrt{\frac{10 \text{ min}}{\tau_{\text{avg}}}}. \quad (2.93)$$

Taken at face value, this suggests that truly quantitative responses—say, to 10% changes in transcription factor concentration—would require hours of integration. This is seldom plausible.

One should not take this rough estimate too literally. I think the message is not the exact value of the limiting precision, but rather that once concentrations fall to the nM range, small changes will be very hard to detect. If cells do detect these small changes, then almost certainly they will be bumping up against the physical limits set by counting molecules, assuming that Berg and Purcell give us a good estimate of these limits. So, this is what we need to check.

Problem 37: Autoregulation. Perhaps the simplest model of transcriptional regulation is one in which a gene regulates its own expression. Let the concentration (or number of molecules) of the protein be g , and assume that n of these molecules bind cooperatively to the promoter region of the gene. If the binding activates expression, and proteins are degraded in a simple first order process with lifetime τ , then it is plausible that the dynamics of g are given by

$$\frac{dg}{dt} = r_{\text{max}} \frac{g^n}{g^n + g_{1/2}^n} - \frac{g}{\tau}. \quad (2.94)$$

(a.) Explain the significance of the parameters r_{max} and $g_{1/2}$. Show that there is a range of these parameters in which the system is bistable. More precisely, show that you can find three steady states, and that two of these are stable and one is unstable.

(b.) Consider an initial condition that is close to one of the stable steady states. What is the time course of relaxation to the steady state? How does the time required for relaxation compare with the lifetime τ of the protein?

(c.) Really the protein binding regulates the synthesis of mRNA, which in turn is translated by the ribosomes into protein. If m is the mRNA concentration (or number of

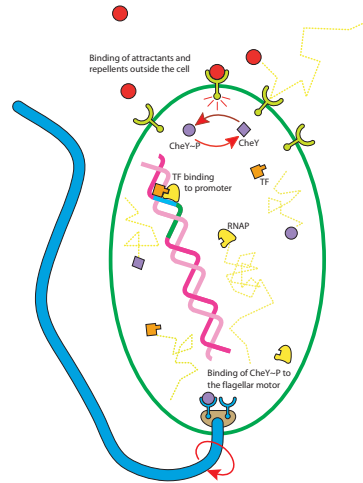


Figure 2.8: Molecule counting in the life of a bacterium. Receptors on the cell surface detect attractants and repellents in the environment, the output of this sensory system is the phosphorylated protein CheY~P, and these molecules are effectively counted by the flagellar motor. In the regulation of gene expression, transcription factors interact with promoter sites on DNA to enhance or repress the activity of the RNA polymerase. With thanks to Sima Setayeshgar for the graphics.

molecules), then a plausible set of equations is

$$\frac{dm}{dt} = e_{\max} \frac{g^n}{g^n + g_{1/2}^n} - \frac{m}{\tau_m} \quad (2.95)$$

$$\frac{dg}{dt} = r_{\text{trans}} m - \frac{g}{\tau_p}, \quad (2.96)$$

where e_{\max} is the maximal transcription (“expression”) rate, r_{trans} is the rate at which mRNA molecules are translated into protein, and the lifetimes of protein and mRNA are τ_p and τ_m , respectively. Under what conditions will this more complete model be well approximated by the simpler model above? Are the steady states of the two models actually different? What about their stability?

(d.) Suppose that instead of activating its own expression, the protein acts as a repressor of its own expression. Find the analog of Eq (2.94) in this case and show that there is only one steady state, and that this state is stable.

(e.) Expand your discussion of the auto-repressor to include the mRNA concentration, as in Eq’s (2.95, 2.96). Find the steady state and linearize the equations around this point. Do you find exponential relaxation toward the steady state for all values of the parameters? Is it possible for the steady state to become unstable? Explain qualitatively what is happening, and go as far as you can in analyzing the situation analytically.

In the lectures, I started by reviewing the fluctuation–dissipation theorem and then using this to derive the spectrum of concentration fluctuations. I think that the review should be an Appendix, and that the simpler problem of concentration fluctuations is indeed a good place to start. What is here is left over from previous years, and I hope still useful. Will go back to revise

In most cases that we know about, biochemical signaling molecules are thought to interact with their receptors through some kinetic process which leads to the establishment of equilibrium between bound and unbound states. If this is the case, we can view the fluctuations in occupancy of a binding site as an example of thermal noise, and we can use the fluctuation–dissipation theorem rather than tracing through the consequences of different microscopic hypotheses about the nature of the interaction between signaling molecules and their targets. We begin with a simple example, to show that we can recover conventional results.

Let's remember what we are trying to accomplish; maybe Fig. 2.8 helps. Bacteria sense chemical signals in their environment, and they do this by having receptors on their surface. In effect the sensory apparatus of the cell “reads” the occupancy of the receptors to infer changes in concentration of the attractants or repellents in the surroundings. But this theme of reading binding site occupancy is repeated over and over again. Essentially every step of biochemical signaling has a similar structure, whether we are talking about the control of the flagellar motor or the control of gene expression. Thus we can hope that understanding noise in receptor occupancy will give us a more general view of the physical limits to cellular signaling.

Consider a binding site for signaling molecules, and let the fractional occupancy of the site be n . If we do not worry about the discreteness of this one site, or about the fluctuations in concentration c of the signaling molecule, we can write a kinetic equation

$$\frac{dn(t)}{dt} = k_+c[1 - n(t)] - k_-n(t). \quad (2.97)$$

This describes the kinetics whereby the system comes to equilibrium, and the free energy F associated with binding is determined by detailed balance,

$$\frac{k_+c}{k_-} = \exp\left(\frac{F}{k_B T}\right). \quad (2.98)$$

If we imagine that thermal fluctuations can lead to small changes in the rate constants, we can linearize Eq. (2.97) to obtain

$$\frac{d\delta n}{dt} = -(k_+c + k_-)\delta n + c(1 - \bar{n})\delta k_+ - \bar{n}\delta k_-. \quad (2.99)$$

But from Eq. (2.98) we have

$$\frac{\delta k_+}{k_+} - \frac{\delta k_-}{k_-} = \frac{\delta F}{k_B T}. \quad (2.100)$$

Applying this constraint to Eq. (2.99) we find that the individual rate constant fluctuations cancel and all that remains is the fluctuation in the thermodynamic binding energy δF :

$$\frac{d\delta n}{dt} = -(k_+c + k_-)\delta n + k_+c(1 - \bar{n})\frac{\delta F}{k_B T}. \quad (2.101)$$

Fourier transforming,

$$\delta n(t) = \int \frac{d\omega}{2\pi} \exp(-i\omega t) \delta \tilde{n}(\omega), \quad (2.102)$$

we can solve Eq. (2.101) to find the frequency dependent susceptibility of the coordinate n to its conjugate force F ,

$$\frac{\delta \tilde{n}(\omega)}{\delta \tilde{F}(\omega)} = \frac{1}{k_B T} \frac{k_+c(1 - \bar{n})}{-i\omega + (k_+c + k_-)}. \quad (2.103)$$

Now we can compute the power spectrum of fluctuations in the occupancy n using the fluctuation–dissipation theorem:

$$\langle \delta n(t) \delta n(t') \rangle = \int \frac{d\omega}{2\pi} \exp[-i\omega(t - t')] S_n(\omega) \quad (2.104)$$

$$S_n(\omega) = \frac{2k_B T}{\omega} \text{Im} \left[\frac{\delta \tilde{n}(\omega)}{\delta \tilde{F}(\omega)} \right] \quad (2.105)$$

$$= \frac{2k_+c(1 - \bar{n})}{\omega^2 + (k_+c + k_-)^2}. \quad (2.106)$$

It is convenient to rewrite this as

$$S_n(\omega) = \langle (\delta n)^2 \rangle \frac{2\tau_c}{1 + (\omega\tau_c)^2}, \quad (2.107)$$

where the total variance is

$$\langle(\delta n)^2\rangle = \int \frac{d\omega}{2\pi} S_n(\omega) = k_B T \left. \frac{\delta \tilde{n}(\omega)}{\delta \tilde{F}(\omega)} \right|_{\omega=0} \quad (2.108)$$

$$= \frac{k_+ c(1 - \bar{n})}{k_+ c + k_-} \quad (2.109)$$

$$= \bar{n}(1 - \bar{n}), \quad (2.110)$$

and the correlation time is given by

$$\tau_c = \frac{1}{k_+ c + k_-}. \quad (2.111)$$

To make sense out of these results, remember what happens if we flip a coin that is biased to produce heads a fraction f of the time. On each trial we count either one or zero heads, so the mean count is f and the mean square count is also f ; the variance is therefore $f(1 - f)$, exactly as in Eq (2.110). The Lorentzian form of the power spectrum is equivalent to an exponential decay of correlations, as expected if the jumping between the two states is a Markov process.

It is interesting that we recover the results for Markovian jumping between two states without making this microscopic model. All we assume is the macroscopic kinetics and that the system is in thermal equilibrium so that we can apply the fluctuation–dissipation theorem. In principle many different microscopic models can describe the molecular phenomena that are at the basis of some observed macroscopic behavior, and we know that many aspects of behavior in thermal equilibrium are independent of these details. The statistics of fluctuations in a chemical kinetic system seems to be an example of this, at least near equilibrium.

The good news, then, is that fluctuations in receptor occupancy are an inevitable consequence of the *macroscopic*, average behavior of receptor–ligand interactions, independent of hypothesis about molecular details. The bad news is that the form of the results doesn't seem very related to the ideas of Berg and Purcell about the precision of concentration measurements. To make these connections clear we need to couple the dynamics of receptor occupancy to the diffusion of the ligand.

When the concentration is allowed to fluctuate we write

$$\frac{dn(t)}{dt} = k_+ c(\vec{x}_0, t)[1 - n(t)] - k_- n(t), \quad (2.112)$$

where the receptor is located at \vec{x}_0 , and

$$\frac{\partial c(\vec{x}, t)}{\partial t} = D \nabla^2 c(\vec{x}, t) - \delta(\vec{x} - \vec{x}_0) \frac{dn(t)}{dt}. \quad (2.113)$$

The first equation is as before, but with notation to remind us that the concentration c is dynamic. The second equation states that the ligand diffuses with diffusion constant D , and when the receptor located at \vec{x}_0 increases its occupancy it removes exactly one molecule from solution at that point.

Following the same steps as above, we find the linear response function

$$\frac{\delta\tilde{n}(\omega)}{\delta\tilde{F}(\omega)} = \frac{k_+c(1-\bar{n})}{k_B T} \frac{1}{-i\omega[1+\Sigma(\omega)]+(k_+\bar{c}+k_-)} \quad (2.114)$$

$$\Sigma(\omega) = k_+(1-\bar{n}) \int \frac{d^3k}{(2\pi)^3} \frac{1}{-i\omega + Dk^2} \quad (2.115)$$

The “self-energy” $\Sigma(\omega)$ is ultraviolet divergent, which can be traced to the delta function in Eq. (2.113); we have assumed that the receptor is infinitely small. A more realistic treatment would give the receptor a finite size, which is equivalent to cutting off the k integrals at some (large) $\Lambda \sim \pi/a$, with a the linear dimension of the receptor. If we imagine mechanisms which read out the receptor occupancy and average over a time τ long compared to the correlation time τ_c of the noise, then the relevant quantity is the low frequency limit of the noise spectrum. Hence,

$$\Sigma(\omega \ll D/a^2) \approx \Sigma(0) = \frac{k_+(1-\bar{n})}{2\pi Da}, \quad (2.116)$$

and

$$\frac{\delta\tilde{n}(\omega)}{\delta\tilde{F}(\omega)} = \frac{k_+\bar{c}(1-\bar{n})}{k_B T} \left[-i\omega \left(1 + \frac{k_+(1-\bar{n})}{2\pi Da} \right) + (k_+\bar{c}+k_-) \right]^{-1}, \quad (2.117)$$

where \bar{c} is the mean concentration. Applying the fluctuation–dissipation theorem once again we find the spectral density of occupancy fluctuations,

$$S_n(\omega) \approx 2k_+\bar{c}(1-\bar{n}) \frac{1+\Sigma(0)}{\omega^2(1+\Sigma(0))^2+(k_+\bar{c}+k_-)^2}. \quad (2.118)$$

We note that the total variance in occupancy is unchanged since this is an equilibrium property of the system while coupling to concentration fluctuations serves only to change the kinetics.

Coupling to concentration fluctuations does serve to renormalize the correlation time of the noise [do you see how to read this conclusion off from Eq (2.118)?],

$$\tau_c \rightarrow \tau_c[1+\Sigma(0)]. \quad (2.119)$$

The new τ_c can be written as

$$\tau_c = \frac{1 - \bar{n}}{k_-} + \frac{\bar{n}(1 - \bar{n})}{2\pi Da\bar{c}}, \quad (2.120)$$

so there is a lower bound on τ_c , independent of the kinetic parameters k_{\pm} ,

$$\tau_c > \frac{\bar{n}(1 - \bar{n})}{2\pi Da\bar{c}}. \quad (2.121)$$

As discussed previously, the relevant quantity is the low frequency limit of the noise spectrum,

$$S_n(\omega = 0) = 2k_+\bar{c}(1 - \bar{n}) \cdot \frac{1 + \Sigma(0)}{(k_+\bar{c} + k_-)^2} \quad (2.122)$$

$$= \frac{2\bar{n}(1 - \bar{n})}{k_+\bar{c} + k_-} + \frac{[\bar{n}(1 - \bar{n})]^2}{\pi Da\bar{c}}. \quad (2.123)$$

If we average for a time τ , then the root-mean-square error in our estimate of n will be

$$\delta n_{\text{rms}} = \sqrt{S_n(0) \cdot \frac{1}{\tau}}, \quad (2.124)$$

and we see that this noise level has a minimum value independent of the kinetic parameters k_{\pm} ,

$$\delta n_{\text{rms}} > \frac{\bar{n}(1 - \bar{n})}{\sqrt{\pi Da\bar{c}\tau}}. \quad (2.125)$$

To relate these results back to the discussion by Berg and Purcell, we note that an overall change in concentration is equivalent to a change in F by an amount equal to the change in chemical potential, so that $\Delta c/\bar{c} \equiv \Delta F/k_B T$. This means that there is an effective spectral density of noise in measuring c given by

$$S_c^{\text{eff}}(\omega) = \left(\frac{\bar{c}}{k_B T} \right)^2 S_F(\omega), \quad (2.126)$$

where the ‘noise force’ spectrum $S_F(\omega)$ is given by the fluctuation–dissipation theorem as

$$S_F(\omega) = \left| \frac{\delta \tilde{n}(\omega)}{\delta \tilde{F}(\omega)} \right|^{-2} S_n(\omega) = -\frac{2k_B T}{\omega} \text{Im} \left[\frac{\delta \tilde{F}(\omega)}{\delta \tilde{n}(\omega)} \right]. \quad (2.127)$$

In the present case we find that

$$S_c^{\text{eff}}(\omega) = \frac{2\bar{c}^2}{k_+\bar{c}(1-\bar{n})} \left[1 + \frac{k_+(1-\bar{n})}{2\pi Da} \right]. \quad (2.128)$$

As before, the accuracy of a measurement which integrates for a time τ is set by

$$\delta c_{\text{rms}} = \sqrt{S_c^{\text{eff}}(0) \cdot \frac{1}{\tau}}, \quad (2.129)$$

and we find again a lower bound which is determined only by the physics of diffusion,

$$\frac{\delta c_{\text{rms}}}{\bar{c}} > \frac{1}{\sqrt{\pi Da\bar{c}\tau}}. \quad (2.130)$$

Note that this is (up to a factor of $\sqrt{\pi}$) exactly the Berg–Purcell result.

In this draft ... I’m going to save you the pain of going through the case of multiple binding sites. In order to understand what happens, however, it is useful to think about a simpler problem. Suppose that we have two balls in a fluid. If they are very far apart, each one experiences a drag force and undergoes Brownian motion, and the Brownian fluctuations in the position of each ball are independent of those in the other. If we bring the two balls together, however, we know that they can influence each other through the fluid: If one ball moves at velocity v_1 it not only experiences a drag force $-\gamma v_1$, it also applies a “coupling” force $\gamma_c(v_1 - v_2)$ to the other ball (which may be moving at velocity v_2 ; clearly if $v_1 = v_2$ there should be no coupling force). If the balls are close enough that γ_c is significant, then in fact the Brownian motions of the two balls become correlated. This correlation can be derived from the fluctuation–dissipation theorem, and it also makes intuitive sense since a random Brownian step of one object applies a force to the other.

If we imagine that positions of the Brownian particles are like receptor occupancies, and an applied force on all the particles is like a change in concentration of the relevant ligand, then diffusion of the ligand serves the same coupling effect as the viscosity of the fluid and should generate correlations among the occupancy fluctuations of nearby receptors. These correlations mean that using the positions or velocities of N Brownian particles to infer the applied force is *not* \sqrt{N} more accurate than using one particle, and similarly using N receptors will not generate a concentration measurement that is \sqrt{N} times more accurate than is obtained with one receptor.

What happens in place of the \sqrt{N} improvement is quite interesting. If we have N receptors, each of size a arrayed on a structure of linear dimension R such as a ring or a sphere, then as N becomes large the limiting concentration noise takes the same form as in Eq (2.130), but the size of the array R takes the place of the receptor size a . If we go back to the intuitive Berg–Purcell argument about counting molecules in a volume and getting a fresh count each time the volume clears from diffusion, what this means is that packing many receptor sites into a region of size R eventually means that we get to count the molecules in a volume $\sim R^3$. There are geometrical factors for different spatial arrangements of the receptors, but like the $\sqrt{\pi}$ in Eq (2.130) these aren't a big deal.

So, at this point we know that the Berg–Purcell argument gives us a *lower bound* on the noise in biochemical signaling. We should now look at real systems to see how close they come to this optimum. This needs a fully digested analysis. In place of this ideal, a collection of references. Of course, start with Berg and Purcell themselves to review the situation in bacterial chemotaxis.

- Cells somewhat larger than *E coli* have the chance to detect gradients across their length. For a discussion of the physical limits in this case, plus a survey of the relevant experiments: Accuracy of direct gradient sensing by single cells. RG Endres & NS Wingreen, *Proc Nat'l Acad Sci (USA)* **105**, 15749–15754 (2008).
- When the brain is developing, neurons have to find their way to their targets in order to make synaptic connections; this too involves a form of chemotaxis. This seems to be extraordinarily sensitive: A new chemotaxis assay shows the extreme sensitivity of axons to molecular gradients. WJ Rosoff et al, *Nature Neurosci* **7**, 678–682 (2004).
- The experimental study of noise in transcriptional regulation was given a huge stimulus by: Stochastic gene expression in a single cell. MB Elowitz, AJ Levine, ED Siggia & PD Swain, *Science* **297**, 1183–1186 (2002).
- For an explicit comparison of precision in transcriptional control with the Berg–Purcell limit: Probing the limits to positional information. T Gregor, DW Tank, EF Wieschaus & W Bialek, *Cell* **130**, 153–164 (2007). See also The role of input noise in transcriptional regulation. G Tkačik, T Gregor & W Bialek, *PLoS One* **3**, e2774 (2008); arXiv:q-bio.MN/0701002 (2007).

At this point I should really give an account of other approaches to noise in chemical systems. We have seen the Langevin method in earlier sections. There is also a master equation approach. For many people the master equation *is* the analysis of chemical noise, but often this path leads away from a proper treatment of diffusion. In place of a digested account,

Problem 38: The simplest master equations. Consider a molecule which is synthesized at some rate s . Neglecting noise or the discreteness of the individual molecules, the number of molecules should change in time as $dN/dt = s$. If we ask about the probability of finding exactly N molecules in the system at time t , this probability $P(N; t)$ obeys the “master equation”

$$\frac{\partial P(N; t)}{\partial t} = sP(N - 1; t) - sP(N; t), \quad (2.131)$$

except of course at $N = 0$ where we have $\partial P(0; t)/\partial t = -sP(0; t)$.

- a. Suppose that we start with no molecules, so that $P(0; t = 0) = 1$ and all other $P(N \neq 0; t = 0) = 0$. We expect that the mean number of molecules after time t will be $\langle N \rangle = st$, and after some thought you might even guess that the distribution of N will be Poisson at each instant of time. Is this correct?
- b. Consider the case where the molecule also is degraded with a lifetime τ , so that

$$\frac{dN}{dt} = s - \frac{N}{\tau}. \quad (2.132)$$

Explain why the master equation is now

$$\frac{\partial P(N; t)}{\partial t} = sP(N - 1; t) - \left(s + \frac{N}{\tau} \right) P(N; t) + \frac{N + 1}{\tau} P(N + 1; t). \quad (2.133)$$

- c. Now that we include decay, we expect that Eq (2.133) has a steady state solution, presumably with a mean number of molecules given by $\bar{N} = s\tau$. Try to find the steady state $P(N)$ analytically. This could be difficult in general but you should be able to find simpler approximations in the limits $\bar{N} \gg 1$ and $\bar{N} \ll 1$.
- d. More generally, if N is not too small we expect that $P(N; t)$ and $P(N \pm 1; t)$ are not too different. Thus we should be able to approximate using a Taylor series,

$$P(N \pm 1; t) \approx P(N; t) \pm \frac{\partial P(N; t)}{\partial N} + \frac{1}{2} \frac{\partial^2 P(N; t)}{\partial N^2}. \quad (2.134)$$

Show that this approximation turns the master equation into something that looks more like the diffusion equation. What is the effective potential in which the “coordinate” N is diffusing? Does this make sense in light of your results in [c]?

- e. Why does it make sense to stop the Taylor series after two derivatives? What happens if we stop after one?

So far we have discussed noise as a small fluctuation around the mean. In a sense, this treats noise as a perturbation. It is also possible that, in the same way that thermal noise can result in a nonzero rate for chemical reactions, noise in chemical kinetics can generate spontaneous switching among otherwise stable states. Much has been written about this. I am less certain that we really understand any particular system. There is, however, some elegant math in the theory, so I would like to come back and discuss this.

The following two problems are concerned with a newly discovered bacterium that responds to a chemical signal by emitting light. The bacteria are roughly spherical, with diameter $d \sim 2 \mu\text{m}$, and hence are clearly visible under the microscope. The chemical signal is shown to be a small protein, presumably secreted by other bacteria; the protein diffuses through the extracellular medium with a diffusion constant $D \sim 10 \mu\text{m}^2/\text{s}$. Very careful experiments establish that each individual bacterium either emits light at full intensity or is essentially dark, and that changing the concentration c of the signaling protein changes the probability of being in the two states. Larger values of c correspond to higher probabilities of being in the light emitting state, so that $p_{\text{light}}(c)$ is monotonically increasing.

Problem 39: Extreme sensitivity, but slowly. There is a specific concentration $c = c_{1/2}$ of the signaling protein such that $p_{\text{light}}(c_{1/2}) = p_{\text{dark}}(c_{1/2}) = 0.5$. When poised at $c = c_{1/2}$ the system switches back and forth between the two states spontaneously at a rate of $\sim 1/\text{hour}$. Remarkably, a change in c by just 10% is sufficient to shift the probabilities from $p_{\text{light}} = 0.5$ to $p_{\text{light}} = 0.9$ or $p_{\text{light}} = 0.1$ when the concentration is increased or decreased, respectively.

(a.) After some confusion in early experiments, it is found that everything said above is true, but the half-maximal concentration $c_{1/2} = 10^{-12}$ M. Is this possible? Justify your answer clearly and quantitatively.

(b.) One group proposes that this extreme sensitivity is not at all surprising, since after all proteins can bind to other proteins with dissociation constants as small as $K_D \sim 10^{-15}$ M. Does this observation of very tight binding have anything to do with the physical limits on sensitivity? Why or why not?

(c.) Another group notes that 10^{-12} M corresponds to $\sim 10^{-3}$ molecules in the volume of the bacterium. They argue that this provides evidence for homeopathy, in which drugs are claimed to retain their effectiveness at extreme dilution, perhaps even to the point where the doses contain less than one molecule on average. Confused as always about these matters, *Nature* publishes a commentary from this group. Can you resolve their confusion?

Problem 40: How simple can it be? Further studies of this new light emitting bacterium aim at identifying the molecules involved. The first such experiment shows that if you block protein synthesis, the system cannot switch between the dark and light states, indicating that the switch involves a change in gene expression rather than (for example)

a change in phosphorylation or methylation states of existing proteins as in chemotaxis. A systematic search which knocks out individual genes, looking for effects on the behavior, finds only one gene that codes for a DNA-binding protein. When this gene is knocked out, all bacteria are permanently dark. More detailed experiments show that these bacteria not only are dark, they actually are not expressing the proteins required for generating light.

(a.) Draw the simplest schematic model suggested by these results. Be sure that your model explains why there are two states (light and dark) rather than a continuum of intermediates, and that your model is consistent with the knock out experiments.

(b.) Assume that the signaling protein binds to some receptor on the surface of the cell and that this triggers a cascade of biochemical events. For simplicity you can imagine that the output of this cascade is some molecule, the concentration of which is proportional to the average occupancy of the receptors over some window of time. Explain how this molecule can couple to your model in [a] to influence the probability of the cell being in the dark or light states.

(c.) Formalize your models from [a] and [b] by writing differential equations for the concentrations of all the relevant species. Show how these equations imply the existence of discrete light/dark states. Can you see directly from the equations why changing the receptor occupancy will shift the balance between these states? It might be hard to explain the behavior near the midpoint ($c = c_{1/2}$), but it should be possible to explain the dominance of the dark state as $c \rightarrow 0$ and the light state as $c \rightarrow \infty$.

(d.) Describe qualitatively all the sources of noise that could enter your model. Do you have any guidance from experiment about which sources are dominant?

(e.) Consider the point where $c = c_{1/2}$. Explain qualitatively what features of your model are responsible for determining the ~ 1 hour time scale for jumping back and forth between the light and dark states.

(f.) See how far you can go in turning your remarks in [e] into an honest calculation!
