MCB137L/237L: Physical Biology of the Cell
Spring 2024
Homework 13 — Extra Credit
(Due 5/7/24 at 2:00pm)

Hernan G. Garcia

This homework draws from biological phenomena and physical models we explored throughout the semester. Note that this homework is for extra credit and that it is completely optional. As we are already going to drop your two lowest-scoring homework, if you complete this homework, we will use it to replace your third lowest-scoring homework. However, we will only do so if it actually *improves* your grade.

1 A minimal genetic switch

In class, we introduced how genetic switches can be constructed using two repressors that repress each other's gene expression. Here, we consider a simpler regulatory architecture that can also result in a genetic switch. Specifically, we will model the self-activation of an activator molecule. For this problem, you might find it useful, to review the phase portrait concept we described in class for the case of mRNA production and degradation, as well as the phase diagram of the logistic equation you had to draw earlier on in the semester.

Figure 1(A) presents a regulatory architecture, where an activator activates its own production. Figure 1(B) shows the states and weights for our model. In this problem, we will ignore mRNA and associate the rate indicated in the figure with the rate of protein production. Here, a promoter has two activator binding sites in its vicinity. In the absence of activators, or in the presence of only one activator, the rate of protein production is r_0. When both activators are bound, the rate is r. The activators bind with a dissociation constant K_d and interact with a cooperativity factor $\omega = e^{-\beta\varepsilon_{int}}$, where ε_{int} is the interaction energy between activators. A is the concentration of activator.

(a) Write down an equation describing the temporal evolution of the number of activators. Consider the rate of production stemming from the model shown in Figure 1, as well as a rate of protein degradation γ. Hint: Remember that the overall rate of production $\langle r \rangle$ of a

system is given by

$$\langle r \rangle = \sum_i p_i r_i, \tag{1}$$

where p_i is the probability of the system being in state i, and r_i is the rate of production when the system is in that state.

(b) Plot the phase diagram for this equation in order to find how many equilibria the system can support. Namely, plot the rates of production and degradation as a function of the activator concentration. Use $K_d = 5$ nM, $\gamma = 0.1$/min, $r_0 = 0.01$ nM/min, and $r = 0.5$ nM/min. Make plots for $\omega = 1$ and $\omega = 10$.

(c) Draw vectors indicating the direction of the concentration change under your plots as we did in class for the mRNA production and degradation case, and as you explored in the homework in the context of the logistic equation. How many equilibrium points do you find? Indicate whether these points correspond to stable or unstable equilibria. You can review the concept of phase portraits by reading "Computational Exploration: Growth Curves and the Logistic Equation" on page 103 of PBoC2, paying special attention to vectors drawn on the lower part of Figure 3.10.

(d) Solve the equation you derived in (a) for different initial conditions, and plot all of them on the same graph. Choose initial conditions that help illustrate how the system can converge to different levels of activator in steady state.

2 Uncovering phase separation in P-granule formation

In the paper by Brangwynne *et al.* (provided on the course website), the authors consider two mechanisms for the accumulation of P granules in the posterior end of the *C. elegans* embryo. First, they posit that P granules could migrate from the anterior end to the posterior end of the embryo. Second, they propose that anterior P granules could be preferentially disassembled or degraded.

(a) Read their paper and write a one-paragraph summary of it. Make sure to explain the various hypotheses they considered and how they tested them.

(b) In their Figure 4, they propose that, upon dissolution in the anterior end, the proteins that make up the P granules diffuse toward the posterior end to take part in granule formation at that location. Assume that these proteins have a reasonable diffusion constant, and estimate the time it takes for these molecules to diffuse throughout the embryo. How do these time scales compare to the overall rates of P granule formation?

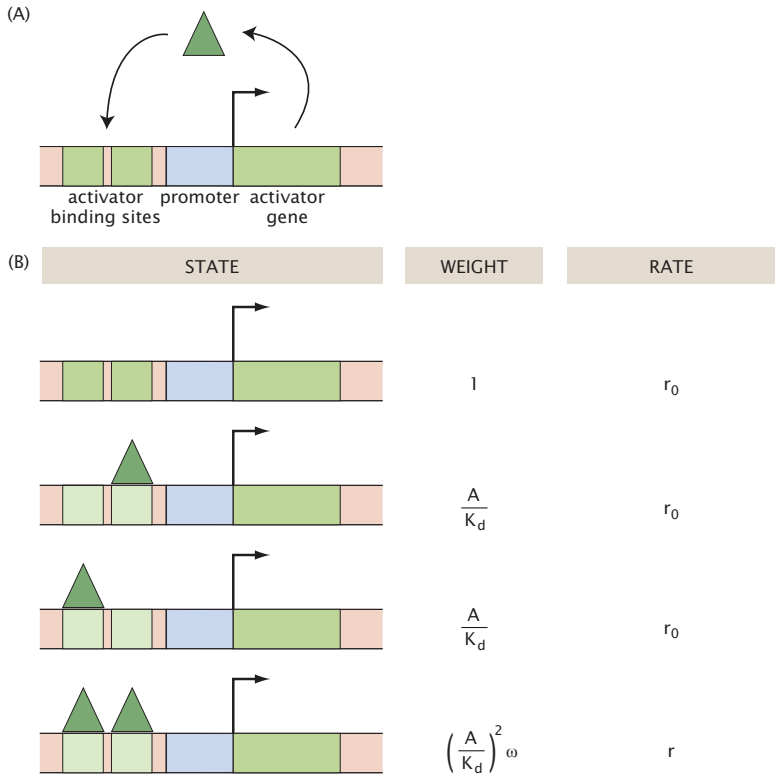

Figure 1: A simple autoactivation switch model. (A) Cartoon of the autoactivation switch. (B) States, weights and rates for the autoactivating genetic switch model.

3 Random walks and biological polymers

Physicists know how to solve just a handful of problems. Fortunately, many dissimilar phenomena in physics and biology alike can be mapped onto such problems for which we know a solution. Here, we explore the mathematical connection between diffusion and the spatial arrangement of polymers such as DNA, actin, and microtubules.

(a) Read the introduction to Section 8.2 of PBoC ("Random Walk Models of Macromolecules View Them as Rigid Segments Connected by Hinges") to learn more about how polymers can be thought of as chains of connected rigid segments. Pay close attention to Figures 8.1 and 8.2. Here, the Kuhn length a is defined as the length of the segments. Look up the Kuhn length for DNA, actin, and microtubules in order to get a feeling for these polymers. Note that you might find reference to the persistence length $\xi_p = a/2$ instead of the Kuhn length.

(b) Now, think of a polymer chain of N segments in 1D. As shown in Figure 8.3 of PBoC each segment can either be pointing to the right of to the left. Given n_R and n_L segments pointing to the right and left, respectively, the position of the end of the chain is given by $L = (n_R - n_L)\, a$. Map this problem onto the diffusion problem we solved in class where we calculated the $\langle x \rangle$ and $\langle x^2 \rangle$ of a random walker that start at the origin shown in Figure 2.

To make this possible, note that each segment can be randomly pointing to the left or right. In particular, calculate $\langle n_R - n_L \rangle$ and $\langle (n_R - n_L)^2 \rangle$ and show that the size of the polymer is given by

$$\text{size} \approx \sqrt{\langle L^2 \rangle} = a\sqrt{N} \tag{2}$$

by repeating graphical the derivation we did in class.

(c) Think of the size of the polymer you derived in (b) as the linear dimension of the blob the polymer will make on a surface such as shown in the figures below. Use the derived formula to estimate the genome length (in μm and bp) of the bacteriophage T2 shown in Figure 1.16 of PBoC and of the *E. coli* in Figure 8.5 of PBoC. How well did your estimate do?

All relevant figures from PBoC can also be found in Figures 3 and 4 below.

4 Saturation of mutant libraries

One of the most important aspect of genetic screens (and life in general) is to recognize when you've reached the point of diminishing returns. To explore this in the context of the genetic screen by Wieschaus and Nüsslein-Volhard, do problem problem 4.5 form PBoC shown in Figure 5. Note that Problem 4.4 mentioned to in the statement refers to Problem 3 of Homework 10. Figure 4.21 from PBoC is shown in Figure 6 while Figure 4.27 is shown in Figure 7.

5 The French flag model

One of the most important and interesting ideas to come out of the idea of positional information contained in morphogen gradients was the so-called French flag model which we will explore here. This model posits that the Bicoid concentration dictates the position of the cephalic furrow. As seen in Figure 8, the idea of the model is that boundaries in the embryo are determined by threshold values of the morphogen. The model predicts that, if the gene dosage of the morphogen gets changed, as seen in the mutant profile, the boundary will still occur at the same value of the morphogen. That hypothesis is enough to determine the shift in boundary position with gene dosage.

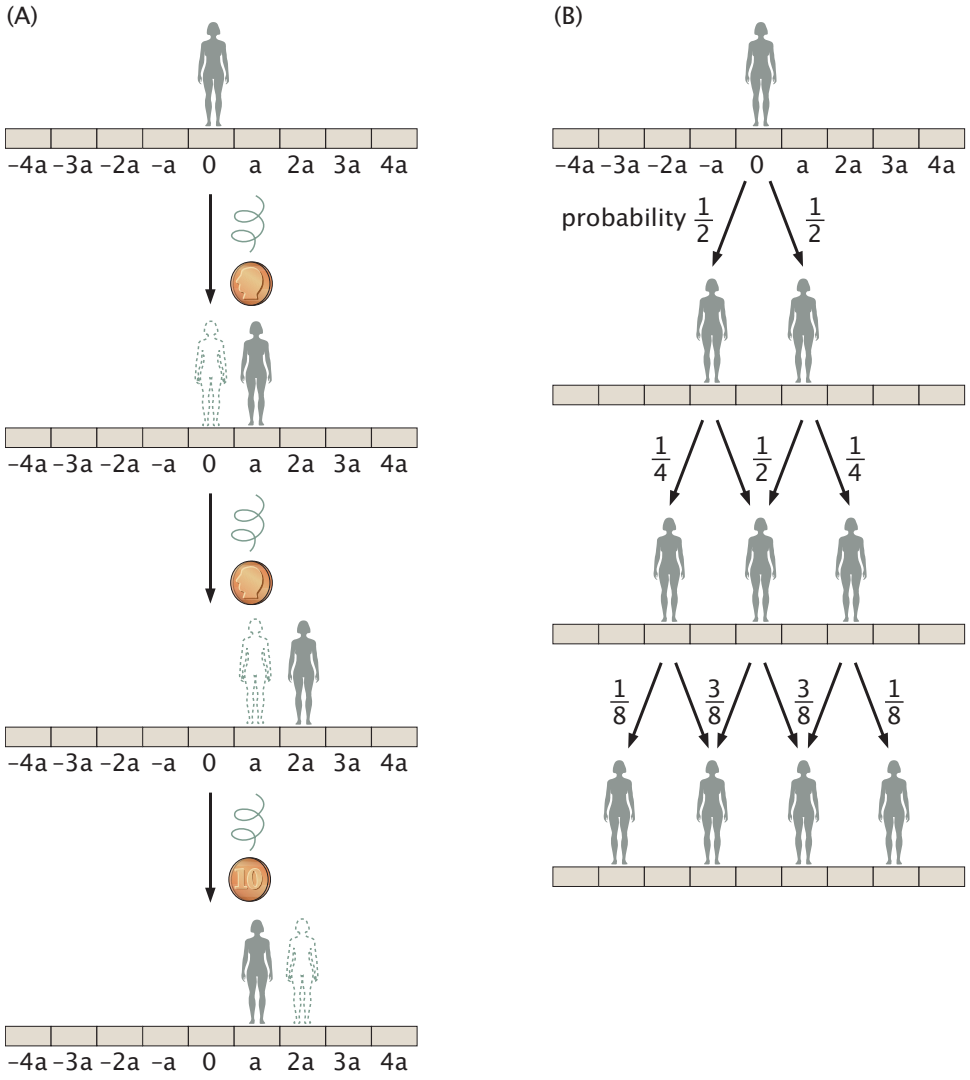

Figure 2: Coin flips and diffusion. (A) Stochastic "simulation" of a coin flipping process with the random walker stepping to the right when a heads is flipped and stepping to the left when a tails is flipped. (B) Scheme for calculating the probability of each and every possible outcome after a total of N steps.

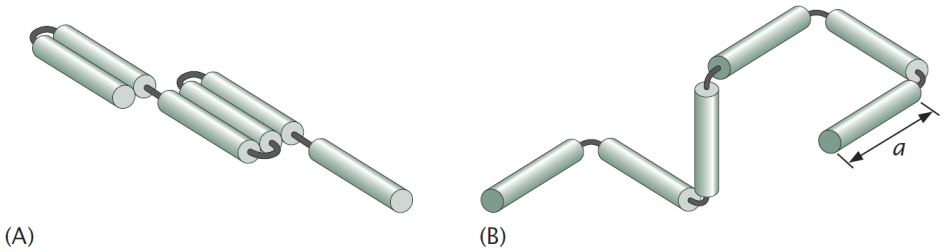

Figure 8.1: Random walk model of a polymer. Schematic representation of (A) a one-dimensional random walk and (B) a three-dimensional random walk as an arrangement of linked segments of length *a*.

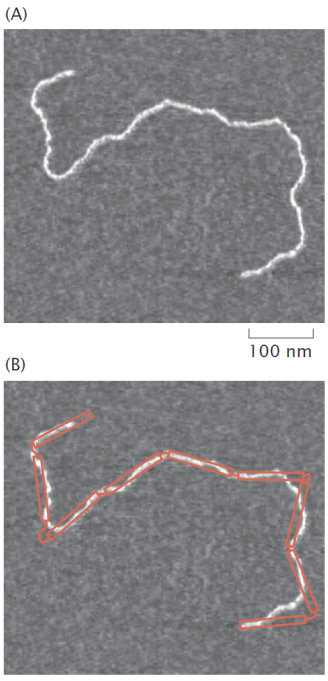

Figure 8.2: DNA as a random walk. (A) Structure of DNA on a surface as seen experimentally using atomic-force microscopy. (B) Representation of the DNA on a surface as a random walk. (Adapted from P. A. Wiggins et al., *Nat. Nanotech.* 1:37, 2006.)

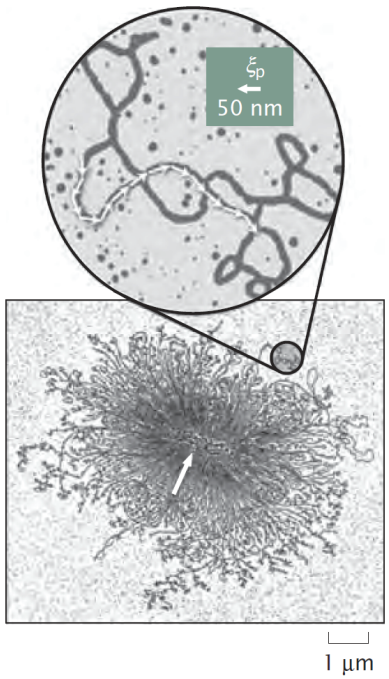

Figure 8.5: Illustration of the spatial extent of a bacterial genome that has escaped the bacterial cell. The expanded region in the figure shows a small segment of the DNA and has a series of arrows on the DNA, each of which has a length equal to the persistence length in order to give a sense of the scale over which the DNA is stiff. (Adapted from an original by Ruth Kavenoff.)

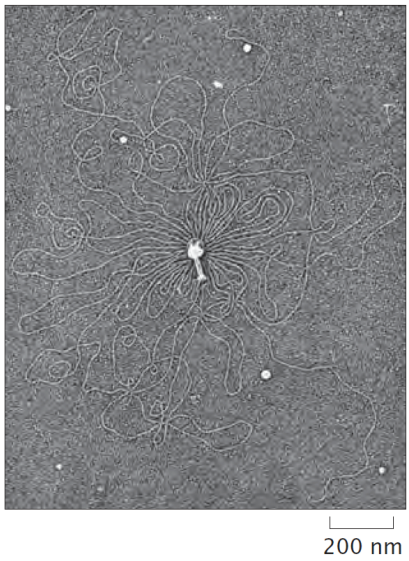

Figure 1.16: Electron microscopy image of a bacteriophage genome that has escaped its capsid. Simple arguments from polymer physics can be used to estimate the genomic size of the DNA by examining the physical size of the randomly spread DNA. We will perform these kinds of calculations in Chapter 8. (Adapted from G. Stent, Molecular Biology of Bacterial Viruses. W. H. Freeman, 1963.)

Figure 3: Figures 8.1, 8.2, 8.5 and 1.16 from PBoC.

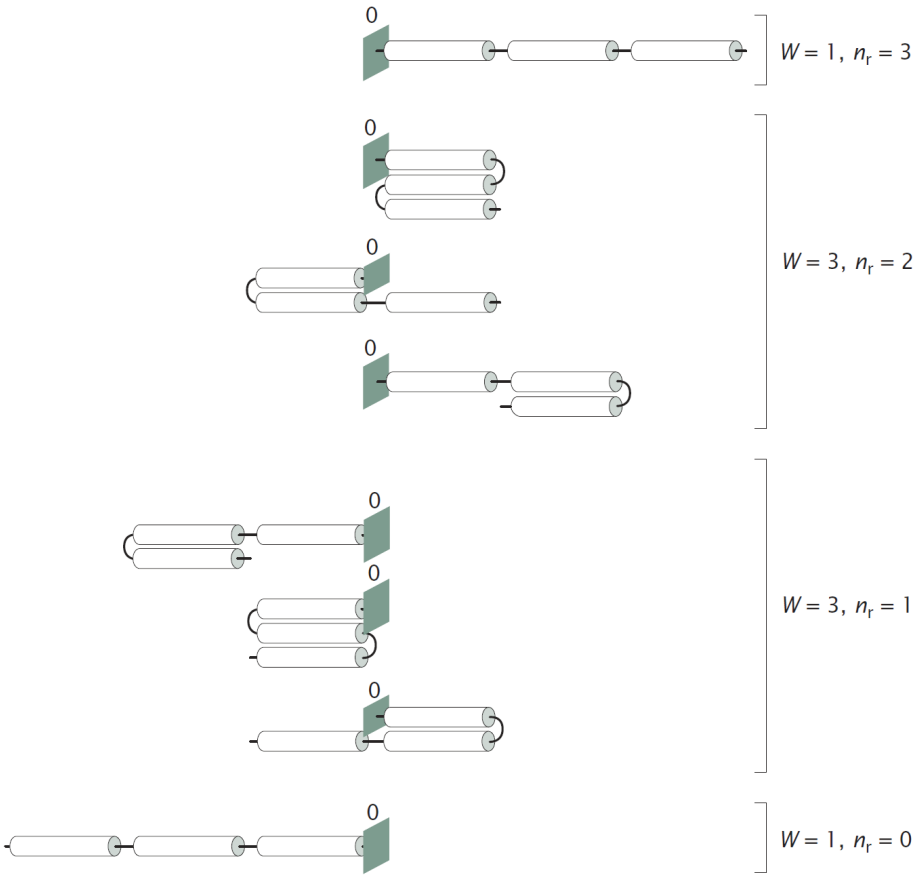

Figure 8.3: Random walk configurations. The schematic shows all of the allowed conformations of a polymer made up of three segments ($2^3 = 8$ conformations) and their corresponding degeneracies.

Figure 4: Figure 8.3 from PBoC.

4.5 Saturation of mutants in libraries

In a set of classic experiments, the second chromosome of *D. melanogaster* was mutagenized and the effects of these mutations characterized based on their phenotype in embryonic development. The experimenters found 272 mutants with phenotypes visibly different from wild-type embryos. However, when they determined the location of the mutations using the method outlined in Figure 4.21 and worked out in Problem 4.4, they discovered that these mutations only mapped to 61 different positions or loci on that chromosome. Figure 4.27 shows how, as more mutants were scored, ever more mutants corresponded to previously identified loci. Using a model that assumes a uniform probability of mutation in any locus, calculate the number of new loci found as a function of the number of mutants isolated. Explain the saturation effect and plot your results against the data. Provide a judgment on whether it is useful to continue searching for loci. (*Hint*: Start by writing down the probability that a specific locus has not been mapped after scoring the first M mutants). Relevant data for this problem are provided on the book's website.

Figure 5: Problem 4.5 from PBoC2.

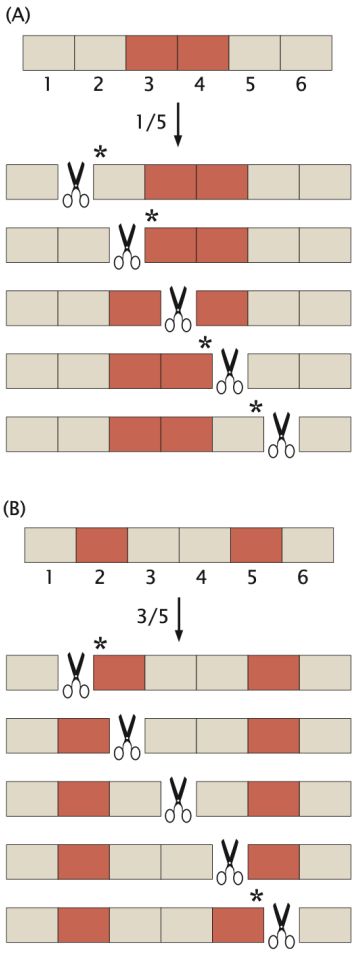

Figure 4.21: Concept of mutation correlation and physical proximity on the gene illustrated by labeling a string in two distinct points and then making random cuts of the string. The probability that the two labels will remain on the same part of the string *after* the cut depends upon their physical proximity *before* the cut. (A) Two mutations (red boxes) that are close to one another are likely to remain on the same part of the string (four times out of five). (B) Two mutations that are further apart are more likely to be separated (remaining together only two times out of five).

Figure 6: Figure 4.21 from PBoC2.

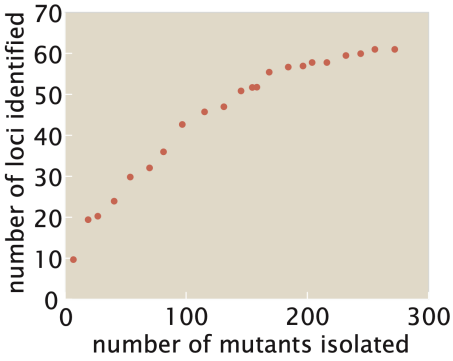

Figure 4.27: Saturation of a mutant library. Number of different identified loci as a function of the number of mutants isolated. (Adapted from C. Nusslein–Volhard et al., *Roux's Arch. Dev. Biol.* 193:267, 1984.)

Figure 7: Figure 4.27 from PBoC2.

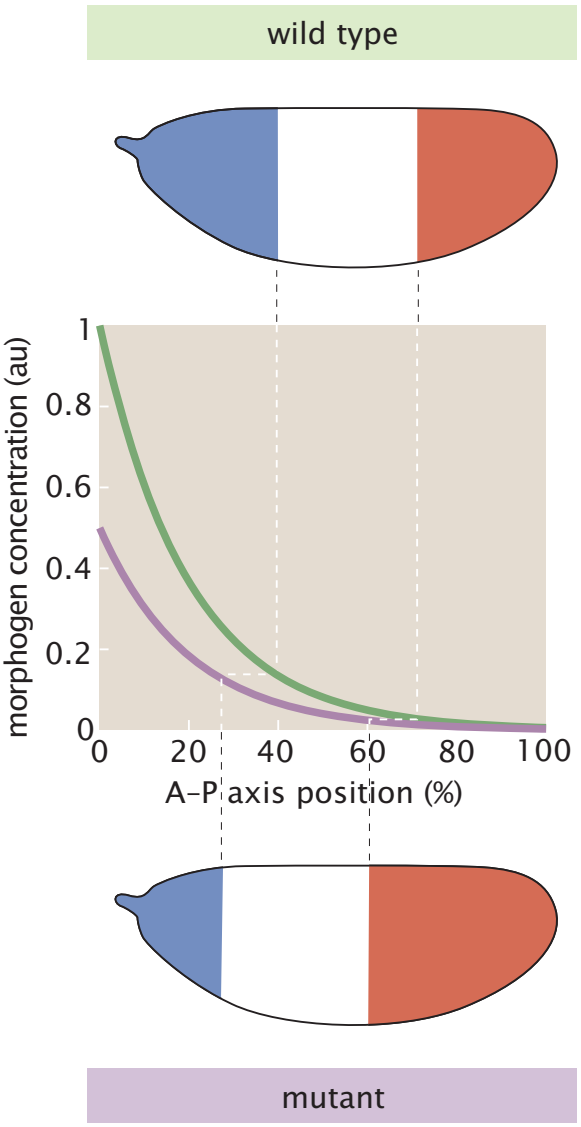

Figure 8: Concept of the French flag model.

To test this model, we will analyze several experiments (Nusslein-Vohlhard and Driever. and Liu *et al.*) where they measured cephalic furrow position as a function of different dosages of the *bicoid* gene in embryos. An exponential gradient of Bicoid is described by

$$Bcd(x, \lambda, \alpha, Bcd_0) = Bcd_0\, \alpha\, e^{-x/\lambda}, \tag{3}$$

where x is the position along the embryo, Bcd_0 is the Bicoid concentration at $x = 0$, λ is the decay constant of the gradient and α is the Bicoid dosage, with $\alpha = 1$ corresponding to the wild-type.

(a) Work out a model that predicts the position of the cephalic furrow x_{new} as a function of the gene dosage α, the morphogen gradient decay length λ and the position of the wild-type cephalic furrow, x_{CF}.

(b) Note that, given a measured $x_{CF} \approx 32\%$ of the embryo length, your model has no free parameters. Compare the prediction from your model with the data for x_{new} vs. α obtained by Nusslein-Vohlhard, and by Driever and Liu *et al.* (provided on the course website). Comment on how well your prediction matches the data that is provided with the homework. What could be going on?