

CS341 SYSTEMS BIOLOGY I

LECTURE 3

1. INTRODUCTION TO KINETIC EQUATIONS

Systems, like the clock have both transcriptional regulation and protein-protein interaction. I am going to begin with protein-protein interactions.

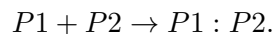
Our goal is to simulate a mathematical model of a biological system based on our understanding of the molecular mechanisms. The molecular mechanisms are comprised of macro molecules (large molecules) such as proteins interacting with each other. One common interaction between proteins is dimerization – when two proteins bind together to form a complex (sometimes called a dimer).

Definition. dimerize – to form a dimer (complex) of two molecules, such as two proteins. We also say the two proteins *bind*. A homodimer is a dimer comprised of two identical proteins. A heterodimer is a dimer comprised of two non-identical proteins.

Today, we will examine this particular interaction in depth, because it forms the basis for much of the modeling we will be reading about and doing later. For example this reaction is key to the function of circadian clocks in fruit flies and mammals.

In biological systems, dimerization is reversible – the dimer may dissociate, producing the two separate proteins. To keep our example simple, we will consider dimerization without the possibility of dissociation,

E.g. if proteins P1 and P2 bind together to form a complex P1:P2, then we think of it as an elementary chemical reaction



Note that this is not an elementary reaction, but we want to study the system at the level of protein interaction. So, at the scale we are interested in, it can be thought of as an elementary reaction. We will use the *language* of chemistry. In fact, we will be borrowing much of our terminology and rules from the study of chemical kinetics.

Definition. chemical kinetics – the study of rates of chemical processes

Typically, in a system, like the clock, there are lots and lots of P1's and P2's in each cell and they are moving around. When they encounter each other, they dimerize.

[Show animation of sparse box. Observe that the molecules bump into each other more frequently when there are more present. Explain the bottom subfigures].

When the concentration is high enough, and the numbers are large enough, we can collect data that is clean enough to determine the rate of reaction.

[Show animation of dense box. Observe that we see the same trends, but they are becoming cleaner. In fact, the bottom right plot is clearly a line. We can calculate its slope. This plot illustrates one of the central principles of chemical kinetics – the law of mass action].

Definition. Law of mass action – the rate of a reaction is proportional the product of the concentrations of the reactants.

E.g. if proteins P1 and P2 bind together to form a complex P1P2, then the rate at which the complex will be formed is proportional to the concentrations of P1 and P2.

Note: In general, square brackets indicate concentration. We will be using them today, but in the future, will leave them off for the sake of simplicity. We don't like cluttered notation!

Under the assumptions that the reactions are occurring continuously and that the system is well-stirred (i.e. that the molecules are arranged uniformly in the space), we write this as a continuous reaction rate equation. The rate at which the concentration of P1P2 (denoted [P1P2]) increases with respect to time is the product of a rate constant k , [P1], and [P2], i.e.

$$\frac{d[P1P2]}{dt} = k[P1][P2]$$

Two important things to note:

- (1) This is a differential equation. We will get back to that in a minute.
- (2) We can calculate k from the bottom right plot. (NOTE: I want to note, however, that this is not how the rate would be measured experimentally. Experimentalists would perform the reaction multiple times, starting each reaction with a different concentration of the reactants, and then measure the concentrations after the reaction has reached equilibrium. And saying even that is not enough because they would be doing this in the presence of a reverse reaction. What I am focussing on here is that in this simulation, we can easily compute the rate constant.) The bottom right plot shows $\Delta[P1P2]/\Delta t$ as a function of $[P1][P2]$. The y-intercept is clearly zero. Calculating the slope provides the rate k :

$$\frac{\Delta[P1P2]}{\Delta t} = k[P1][P2]$$

and since

$$\frac{d[P1P2]}{dt} \sim \frac{\Delta[P1P2]}{\Delta t}$$

we can use this data to provide the rate for the differential equation which describes the system.

2. REACTION RATE EQUATIONS

To model the entire system, we define a rate equation for each chemical species (in our case, for each type of protein).

The general form is like this

$$\frac{dP}{dt} = \text{rate_of_production} - \text{rate_of_consumption}$$

So far, we have one reaction, involving three species (P1, P2, and P1P2). We have already seen the rate equation for P1P2 (and note that is complete), but what about P1 and P2?

To write the equation for P1, we ask ourselves, in which reactions is P1 involved? For each reaction, we follow this simple algorithm:

- If P1 consumed by the reaction, subtract the reaction rate
- If P1 is produced by the reaction, add the reaction rate

Our complete model is then

$$\begin{aligned}\frac{d[P1]}{dt} &= -k[P1][P2] \\ \frac{d[P2]}{dt} &= -k[P1][P2] \\ \frac{d[P1P2]}{dt} &= k[P1][P2]\end{aligned}$$

3. REPRESENTING ODES IN VECTORS/ARRAYS

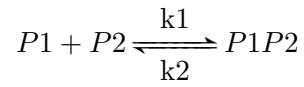
For the dimerization model, we have three equations. We can think of these as three functions. But each of these functions is dependent on the same set of variables. So, it is better to group them together and think of them as a vector equation.

In Python, that means writing a function that takes a vector of state variables as its input and returns a vector (array) of rates as its output. So we are doing vector calculus, but we can think of it as functions that take arrays as input and return arrays as output.

The code is in a Jupyter Notebook.

4. ADDING THE REVERSE REACTION

Dimers can also unbind, so we really should model it as reversible:



where $k1$ is the rate constant for binding and $k2$ is the rate constant for unbinding.

The ODEs are

$$\begin{aligned}\frac{d[P1]}{dt} &= -k1[P1][P2] + k2[P1P2] \\ \frac{d[P2]}{dt} &= -k1[P1][P2] + k2[P1P2] \\ \frac{d[P1P2]}{dt} &= k1[P1][P2] - k2[P1P2]\end{aligned}$$