Introduction to Bioinformatics (2003)
Scenario 4: Metabolic Modeling

Outline:

I. Where you're going vs Where you're going to
II. Chemical reactions
III. Enzymatically catalyzed reactions
IV. Inhibition of enzymatically catalyzed reactions
V. Eisenthal and Cornish-Bowden (1998)

I. The difference between knowing where you're going and knowing where you're going to

You are a defense analyst, specializing in the workings of a secretive Asian country. They have just developed a highly accurate ballistic missile. It didn't take any special spook skills to figure this out, since they triumphantly broadcast the test launch for all the world to see, from liftoff to splashdown precisely on a handkerchief marked with an X. The site of landing in the Indian Ocean was announced beforehand and is precisely known. However, the launching site is a closely guarded secret, one that you intend to uncover.

So there you are, scrutinizing the film of the launch, frame by frame. Of course there are no distinguishing background features that could give you a clue as to the location, but you've found that a close inspection of the contrails emanating from the missile you can determine the instantaneous velocity at every point in the film. You also measure the angle of the missile. Finally, you know that each frame is separated from the next by $1/24^{th}$ of a second. From this information, you hope to recreate the path of the launch and work backwards from the known landing site to predict the location of the secret launch base.

At each point in the film, you multiply the velocity of the missile by $1/24^{th}$ second to determine the distance traveled by the missile to the position shown in the next frame. You repeat this procedure for all the frames to recreate the trajectory (Fig. 1). From this calculated path and the known angle of incidence at the handkerchief, you can trace the missile back to its source, a balloon factory in a small desert community. This enables you to wire them your warmest congratulations.

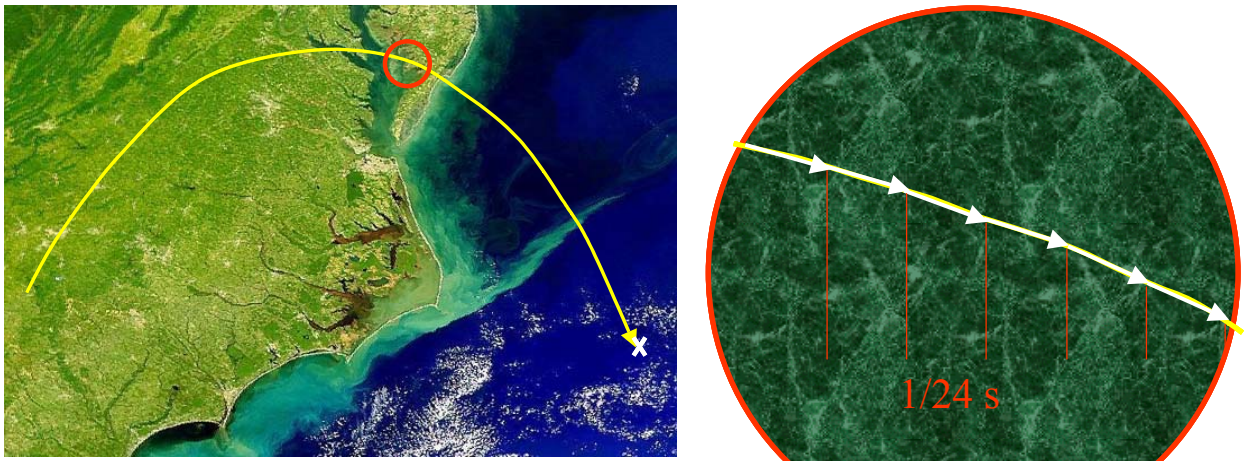

Figure 1: Trajectory of missile test. **A.** Actual path of missile. **B.** Recreated path determined from the observed angle of the missile and from the calculated distance.

However unlikely the above scenario may seem to you, the strategy of recreating a trajectory from instantaneous velocity is not at all uncommon. Consider the centerfielder who catches a glimpse of a fly ball. Does he think, "Let's see, the formula for a parabola is…." No, more likely, he notes the current position and the current velocity (and direction) and predicts the flight of the ball and where he ought to go to catch it.

In certain ideal circumstances (e.g. bodies moving in frictionless space), it is possible to derive a precise expression that predicts location at every moment in time. This is often not possible in real, frictionful life. In such circumstances, the only practical alternative is often to proceed as in the example, by determining the current velocity and applying it to the current location to predict the next location.

II. Chemical Reactions

The nature of chemical reactions often makes it relatively easy to predict the instantaneous velocity but next to impossible to predict an overall equation of location, just as it was in the case of the task of determining the missile's trajectory. First of all, what are the logical counterparts of "velocity" and "location" in the case of chemical reactions? Let's take as an example one of the reactions of glycoloysis:

(1) glucose + ATP → glucose-6-phosphate + ADP

By velocity of the reaction I mean the rate at which product (glucose-6-phosphate; G6P) is being formed. I could also write this as d[G6P]/dt, the rate of change of G6P at a given moment in time. Suppose that this reaction took place at some fixed frequency whenever a molecule of glucose banged into a molecule of ATP. Then certainly there would be more opportunity for the reaction as the concentration of glucose increased and likewise more opportunity if the concentration of ATP increased. In short:

(2) velocity = d[G6P]/dt = **k** [glucose] [ATP]

where **k** is related to the fraction of times a collision leads to a productive reaction. So, if I knew the concentration of glucose and ATP at a given moment in time and the constant **k**, I'd be able to determine the rate at which product is produced.

Let's consider a simpler reaction:

(3) adenosine diphosphate-arsenate → adenosine diphosphate + arsenate

Adenosine diphosphate-arsenate (ADP-As) is formed during glycolysis when arsenate is present, substituting for phosphate in reactions leading to the formation of ATP. ADP-As spontaneously breaks down to ADP + arsenate, and so the metabolic energy captured in the formation of ADP-As is lost, one reason why arsenic is a potent toxin. This reaction if similar to Reaction (1) because there is only one substrate (there is a second reason we'll get to momentarily). The velocity of the reaction should be:

(4) velocity = d[ADP]/dt = **k** [ADP-As]

and since the rate of formation of ADP is the same as the rate of loss of ADP-As:

(5) d[ADP]/dt = - d[ADP-As]/dt = **k** [ADP-As]

So, just as you knew the velocity of the missile at every point, we know from this equation the velocity of the reaction at every concentration of substrate, and we can use this velocity to predict the new concentration of substrate a moment later:

(6) $[\text{ADP-As}]_{to + dt} = [\text{ADP-As}]_{to} + (d[\text{ADP-As}]/dt)\ dt$

That is, the new position is the old position + the velocity times the unit of time. This should make some sense. If you're in a car, your new position is your old position plus your speed x time. If your at a 30-mile marker going 60 mph and you travel 1 hr, your new position is 30mi + (60 mph)(1 hr) = 90mi. From this equation, we should be able to predict the entire course of a reaction.

SQ1. Let's try it. If you go to the Scenario web site, you'll find a link to ADP-As.pl, a program that models Reaction (3). It presumes arbitrarily that the starting concentration of ADP-As is 1 mM and that the rate constant k is 1 sec/mM (I made this up). Download and run the program, giving you a file containing the predicted level of ADP-As over time. You can upload this file into Excel and plot the results.

How good were the predictions? In this very simple case it is possible to analytically predict the expected concentration of ADP-As at any point in time by integration:[1]

(7) $d[\text{ADP-As}]/dt = -\mathbf{k}\,[\text{ADP-As}]$

(8) $d[\text{ADP-As}] / [\text{ADP-As}] = -\mathbf{k}\ dt$ # Multiply through by the denominators

(9) $\ln[\text{ADP-As}] = -\mathbf{k}\,\mathbf{t} \mid \mathbf{t} = \mathbf{t}_0 \,..\, \mathbf{t}$ # Integrate

(10) $\ln[\text{ADP-As}]\mathbf{t} - \ln[\text{ADP-As}]\mathbf{t_o}$
$= \ln([\text{ADP-As}]\mathbf{t} / [\text{ADP-As}]\mathbf{t_o})$
$= -\mathbf{k}\,(\mathbf{t} - \mathbf{t_o})$ # Evaluate between $\mathbf{t_o}$ and $\mathbf{t}$

(11) $[\text{ADP-As}]\mathbf{t} / [\text{ADP-As}]\mathbf{t_o} = e^{-\mathbf{k}(\mathbf{t} - \mathbf{to})}$

(12) $[\text{ADP-As}]\mathbf{t} = e^{-\mathbf{k}\,\mathbf{t}}$

SQ2. Armed with this function, download and run Plot_function (which already has the function put into it), and bring those results into the same Excel page as before. Plot the predicted and calculated levels of ADP-As on the same graph so you can see how close they are.

SQ3. Why is it that the *predicted* levels are all lower than the *calculated levels*?

SQ4. Modify the two programs to make produce numbers that bring the two curves closer together. What's the key trick?

III. Enzymatically catalyzed reactions

Equation (2) seems to imply that the higher the concentration of glucose, the faster the reaction will go. Glucose is pretty cheap… all I need to do is dump in a bag full of sugar, and I should make lots of the expensive chemical G6P. A sure-thing money maker? Don't count on it. Doubling the amount of glucose ***will*** double the rate of the reaction, true, but it will double it from around 1 mmole per 20 million years to 1 per 10 million years. Banging ATP into glucose is not a good way to get glucose-6-phosphate!

---

[1] For those of you not into integration, why not go through it anyway? It's just algebra, except for step 9.

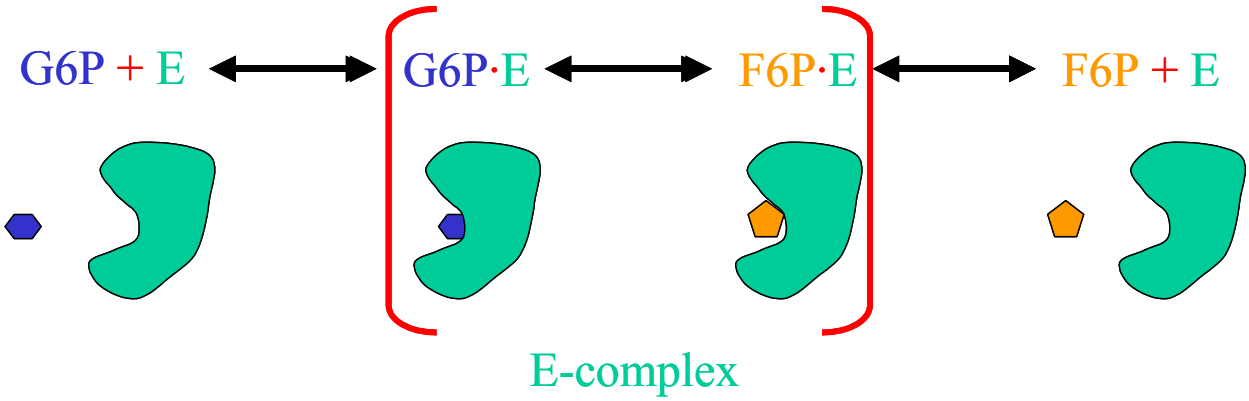

Figure 2: Molecular events in enzymatically catalyzed isomerization of glucose-6-phosphate (G6P) to fructose-6-phosphate (F6P). Blue hexagon represents G6P. Orange pentagon represents F6P. Big green blob represents the enzyme (E) phosphoglucose isomerase.

This is why enzymes exist. Unfortunately, the existence of enzymes also complicates the formulation of rate equations. Figure 2 illustrates the chemical events that occur during a enzyme catalyzed reaction. Instead of chemicals banging into each other, they bang into an enzyme that holds them in place and facilitate a specific transformation. Nonetheless, each part of the series of events is a chemical reaction that can be characterized by a chemical rate equation. So in the example above, the formation of the G6P-enzyme complex occurs at a velocity:

(13) $d[G6P{\cdot}E]/dt = \mathbf{k}_f\ [G6P]\ [E]$

(the rate constant is given as $\mathbf{k}_f$ to indicate that it is the constant for the ***forward*** direction). Likewise, the velocity of the release of G6P from the complex occurs at a velocity:

(14) $d[G6P]/dt = \mathbf{k}_r\ [G6P{\cdot}E]$

($\mathbf{k}_r$ indicating the constant for the reverse reaction). The velocity of the overall reaction is the production of the product F6P, given as:

(15) $\boldsymbol{v} = d[F6P]/dt = \mathbf{k}_c\ [F6P{\cdot}E]$

The concentration of the F6P·enzyme complex cannot be readily determined, but fortunately, the velocity can be expressed in terms of measurable quantities (I'll skip the derivation for now):

(16) $\boldsymbol{v} = d[S]/dt = \mathbf{k_c}\ E_{total}\ (\mathbf{S}\ /\ (\mathbf{S} + K_m) = (\mathbf{S}\ \mathbf{V_{max}})\ /\ (\mathbf{S} + K_m)$

(where **S** is the concentration of substrate G6P, K_m is a collection of rate constants, and $\mathbf{V_{max}}$ is the maximal velocity of the reaction).

Let's examine this equation. When the substrate level is very high, the enzyme has bound as much substrate as it can and the overall reaction is going as fast as it can. Under such circumstances, $\mathbf{S} >> K_m$ and so $\boldsymbol{v} = \mathbf{V_{max}}$. Suppose instead that there's hardly any substrate at all, then $\mathbf{S} << K_m$ and $\boldsymbol{v} = \mathbf{S}\,\mathbf{V_{max}} / K_m$. In other words, the velocity is directly proportional to the substrate concentration, which makes sense, because if the enzyme is almost all free, then doubling the substrate level should double the amount of enzyme that binds substrate. Finally, if $\mathbf{S} = K_m$, then $\boldsymbol{v} = (K_m\,\mathbf{V_{max}}) / K_m + K_m) = \mathbf{V_{max}} / 2$. In other words, K_m is the concentration of substrate at which the velocity is exactly half the maximal velocity.

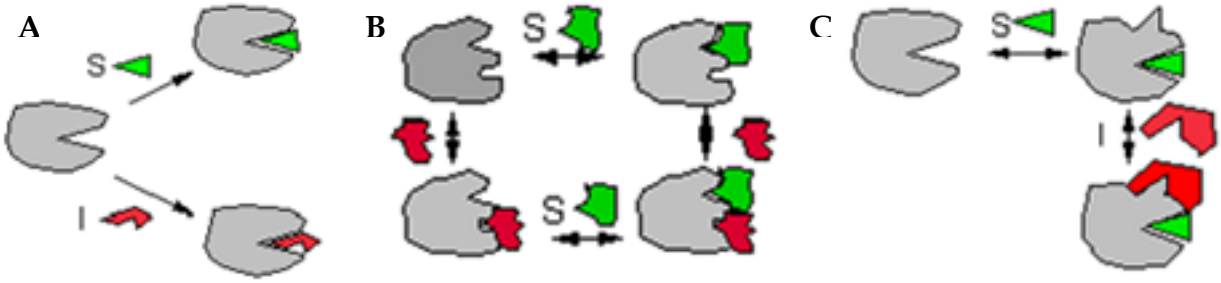

Figure 3: Types of inhibition of enzymatic reactions. Gray blobs represent enzymes, green blobs substrate, and red blobs inhibitors. (**A**) Competitive inhibition. (**B**) Noncompetitive inhibition. (**C**) Uncompetitive inhibition. (Courtesy of Henry Jakubowski, College of St. Benedict/St. Johns University; http://employees.csbsju.edu/hjakubowski/classes/ch331/transkinetics/olinhibition.html)

SQ5. Write an expression that gives the concentration of F6P at time *t* + dt.

SQ6. From these limiting conditions, draw an approximate curve relating S (x-axis) to *v* (y-axis). Indicate also K_m.

IV. Inhibition of enzymatically-catalyzed reactions

The goal of our scenario is to find an inhibitor of some reaction in the glycolytic pathway that most efficiently starves trypanosomes for ATP. In order to fit inhibition into the model, it, like velocity, must be expressed in quantitative terms.

One can imagine a variety of ways that reactions catalyzed by enzymes might be inhibited (Fig. 3). One obvious way is that an inhibitory chemical might look like the substrate and compete with it for binding to the enzyme. This type of inhibition is called competitive inhibition. The hallmark of competitive inhibition is that it is overcome when the substrate concentration is sufficiently higher than the inhibitor concentration. Another type is noncompetitive inhibition, where the inhibitor binds to the enzyme at a different site than the substrate. The binding of the inhibitor prevents the enzyme from performing catalysis. As a result, the effective $\mathbf{V_{max}}$ of the reaction is lowered. The higher the level of inhibitor, the more the $\mathbf{V_{max}}$ is lowered. Finally, there is uncompetitive inhibition, where the inhibitor binds specifically to the substrate-enzyme complex, preventing catalysis. In this case, the binding of inhibitor affects both $\mathbf{V_{max}}$ and K_m.

Uncompetitive and noncompetitive inhibitors are generally more desirable in drugs, because they can have an effect on catalysis even when the substrate is in excess, as is usually the case.

VI. Eisenthal and Cornish-Bowden (1998)

Robert Eisenthal and Athel Cornish-Bowden has performed the very analysis our scenario demanded: modeling glycolysis of *Trypansosoma brucei* and determining which step may be inhibited to the greatest effect. As you might imagine, it is a rather complex article. Download the article and, for now, do the following:

A. Read the **Introduction**, mostly for pleasure.

B. Glance at the **Experimental Model** section, focusing mostly on the parts that are understandable from the foregoing discussion

C. Skim the **Results**, but pay special attention to **Table I** and consider how you would use it to do the modeling yourself.

D. Come tomorrow armed with questions.

SQ7. Consider Eq.2 in the paper (described as a generic equation). Suppose that the reaction proceeds only in the forward direction (say the product is continuously being destroyed, so P_1 and P_2 are both 0). What is the resulting equation?

SQ8. Starting with the equation you found in SQ7, suppose further that there is only one substrate, S_1. The second substrate, S_2, can be considered to be at an inexhaustibly high concentration. This would be the case if the substrate were water, for example. What is the resulting equation now?