# 20.320, notes for 9/25

Last time

We were talking about kinases and their activation loops.
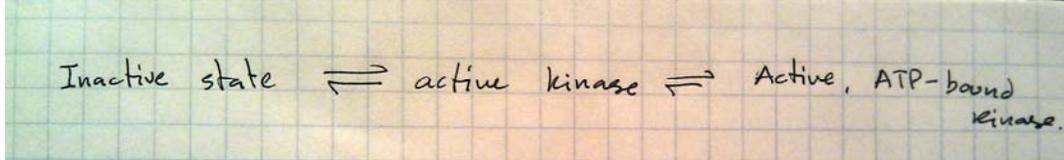
Let's now go into greater detail of how kinases work and how they fit into signaling cascades. We'll also tie it to EGFR.

EGFR normally exists at about 10^4 copies per cell. In cancers, it gets overexpressed up to 10^6. Why does overexpression lead to an increased signal? One of the reasons is simply thermodynamic.

Kinases, like all proteins, have several different conformational states (see above), and those states all having different energy levels. Kinases tend to live in the lowest-energy state, which is the inactive state for them. What phosphorylation does is push the molecules to spend more time in their active state. There, they can be locked in place by ATP and then dimerize and send a signal. Protein populations, as you know, always have a few members that randomly get enough energy from their environment to jump into higher-energy state. A small fraction of the population will be in the active state all the time, though it can't signal because the state is short-lived and you need two proteins like that to dimerize.

What happens when the receptors are overexpressed is that suddenly there's 100 times more, and so there's enough for a signal to be produced just from receptors that happened to randomly shift into the active state (even in the absence of ATP) and bump into each other. That's the reason why you can get increased activation even in the total absence of a signal.

We talked last class about Type II inhibitors, in which the DFG region (the hinge, named for its amino acid sequence Asp-Phe-Gly) is locked in the "DFG-out" state (the active state).

Constitutive Kinases

If a phosphate on the activation loop locks a protein in the active state, how can we mimic this in the lab? Well, remember that the effect of that phosphate is caused by its electrostatic charge. Phosphate groups are very negative. One way to mimic this in lab is to mutate the receptor so that a few of the amino acids in the activation loop are suddenly more negative.

S,T,Y --> mutate to --> D,E

This results in increased receptor activity. We can make these mutations in the lab to create constitutively active kinases. It also happens in nature, where all of a sudden you have kinases with 1000x increased activity. These are common in melanoma and lung cancer.

Modeling kinases

These are the kinetic equations for kinase activity.

$$E_1 + E_2 \underset{K_{-1A}}{\overset{K_{1A}}{\rightleftharpoons}} E_1 E_2 \overset{K_{2A}}{\longrightarrow} E_1^* + E_2 \qquad\qquad E_1^* = \text{activated kinase}$$

$$E_1^* + S \underset{K_{-1B}}{\overset{K_{1B}}{\rightleftharpoons}} E_1^* S \overset{K_{2B}}{\longrightarrow} E_1^* + S - PO_4$$
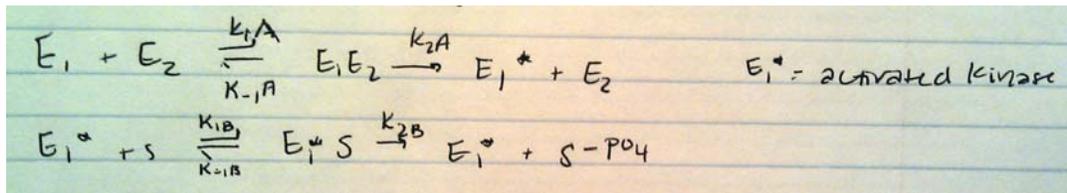
It's tempting to write the equations in this other shorthand form that neglects the formation of an intermediate complex. DO NOT do this, for it neglects the effect of affinity and will mess up your models.

why can't we write $E_1 + E_2 \overset{K_{phos}}{\longrightarrow} E_1^* + E_2$?
- doesn't account for interactions, K_M
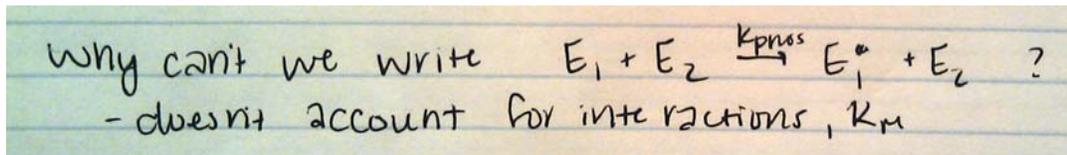
How to shut down kinases.

We must also remember to think about **phosphatases**, the proteins removing the phosphates put in by kinases.

$$E_1^* + P'ase \rightleftharpoons E_1 \cdot p'ase \overset{K_{2P}}{\longrightarrow} E_1 + P'ase$$
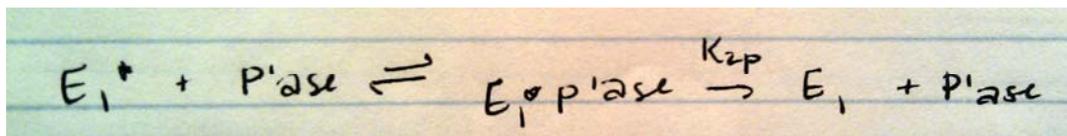
PTEN is a particular phosphatase that is mutated in 40% of cancers. It's a dual-specificity phosphatase that dephosphorylates both lipids and tyrosines. The lipid it dephosphorylates is Phosphoinositol 2,4,5 phosphate (PIP3), which gets de-activated into PIP2. PTEN shuts down the PIP3 signal, which would otherwise be responsible for pro-survival and anti-apoptotic signaling. When PTEN disappears in cancers, the signals endure and you get cancerous proliferation.
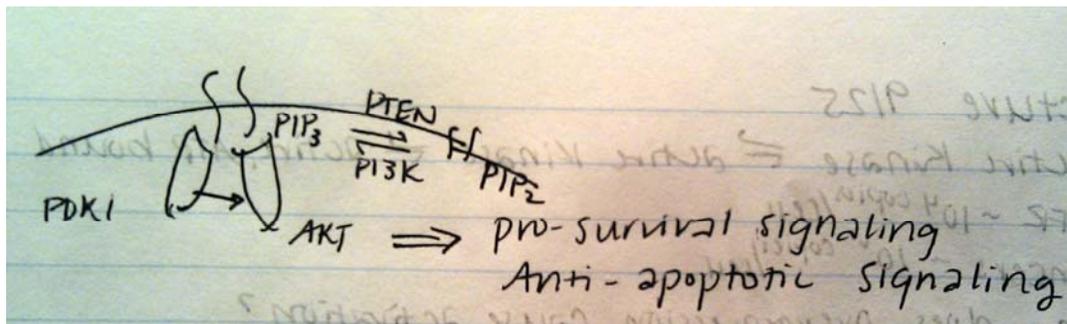
There is also the **Ubiquitin pathway**. Ubiquitin ligases tag proteins with ubiquitin and target them for proteosomal degradation.
- Mutate out the ligase, and you have a problem
- A common mutation is C-Cbl, the ubiquitin ligase that shuts down activated EGFR. Common in multiple cancers.

You can think of most of these regulatory systems as push-pull interactions between activation and deactivation pathways. This is a recurring theme in biology.

Multiple substrates

Kinases usually have 100s of substrates, despite how we tend to talk about them. The models quickly turn very difficult. But what does this do for the kinase? Imagine you had a kinase with two different protein substrates.

Substrate	Km
S_1	10nM
S_2	200nM

Even if you had the same k_{cat} and concentration for both of them, a limited amount of kinase will result in a much higher phosphorylation rate for S_1. This is the effect of different K_m values. This is what happens in the cell. K_m ratios result in ratios of the phosphorylated proteins. Note that if you did this in a tube, then after S_1 was depleted you'd finally see the phosphorylation of S_2. In the cell, where you have dozens of substrates competing like this, the net result is built-in time delays. With no other form of regulation, a single enzyme will produce successive "waves" of substrate phosphorylation.

Modeling Abstraction

Modeling stuff like this gets very difficult very quickly. We will cover ways of simplifying the system in order to model it. Instead of modeling all the complexity at the start, we will start by abstracting first and then adding complexity as needed.

Perhaps the greatest abstraction is the creation of a black box. We can model the cell as a closed system that takes in a cue, creates an internal signal, and produces a response. What are some potential **cues** and **responses**?

Cues
- pH
- Growth factors
- Surface interactions
- Cell-cell interactions
- DNA damage
- Nutrients

Responses
- Apoptosis
- Proliferation
- Migration
- Senescence
- Necrosis

If we were biologists, we'd concern ourselves with the fine details of how proteins drive each of the details of all this. As bioengineers, we'll learn enough about the systems to understand how to abstract them and create useful models of the system.

The MAP Kinase Cascade

Finally, we get to this one. It's a system conserved through all eukarya, from yeast to human. There are several kinase cascades, but MAP Kinase cascades regulate the majority of these cue-signal-response systems. There are several MAP Kinases:

Human MAPK's

- **Extracellular signal-Regulated Kinases**
  - ERK1, ERK2
  - Responsible for proliferation, migration apoptosis, invasion, transcription, etc.
  - They have been linked to almost everything in biology, but their main purpose is to take in stimuli and turn them into proliferation signals.
- **p38**, also known as the Stress Activated Protein Kinase
  - Responds to low nutrients, DNA damage, kinase inhibitors
- **JNK** --> c-Jun n-Terminal Kinase
  - Activated by inflammation, part of the immune system

      ○ Also regulates cells' insulin resistance.

All MAPK cascades look the same. The only thing that changes in different systems and organisms is the particular players that take on any particular role. Huang & Ferrel, PNAS, 1996, are a great example of modeling the MAPK cascade. It will be posted online.

So what does the MAPK cascade look like?

NOTE: The asterisk indicates activated. This is usually synonymous with phosphorylated, even if in some systems this activation may not be caused by phosphorylation. The model still works.
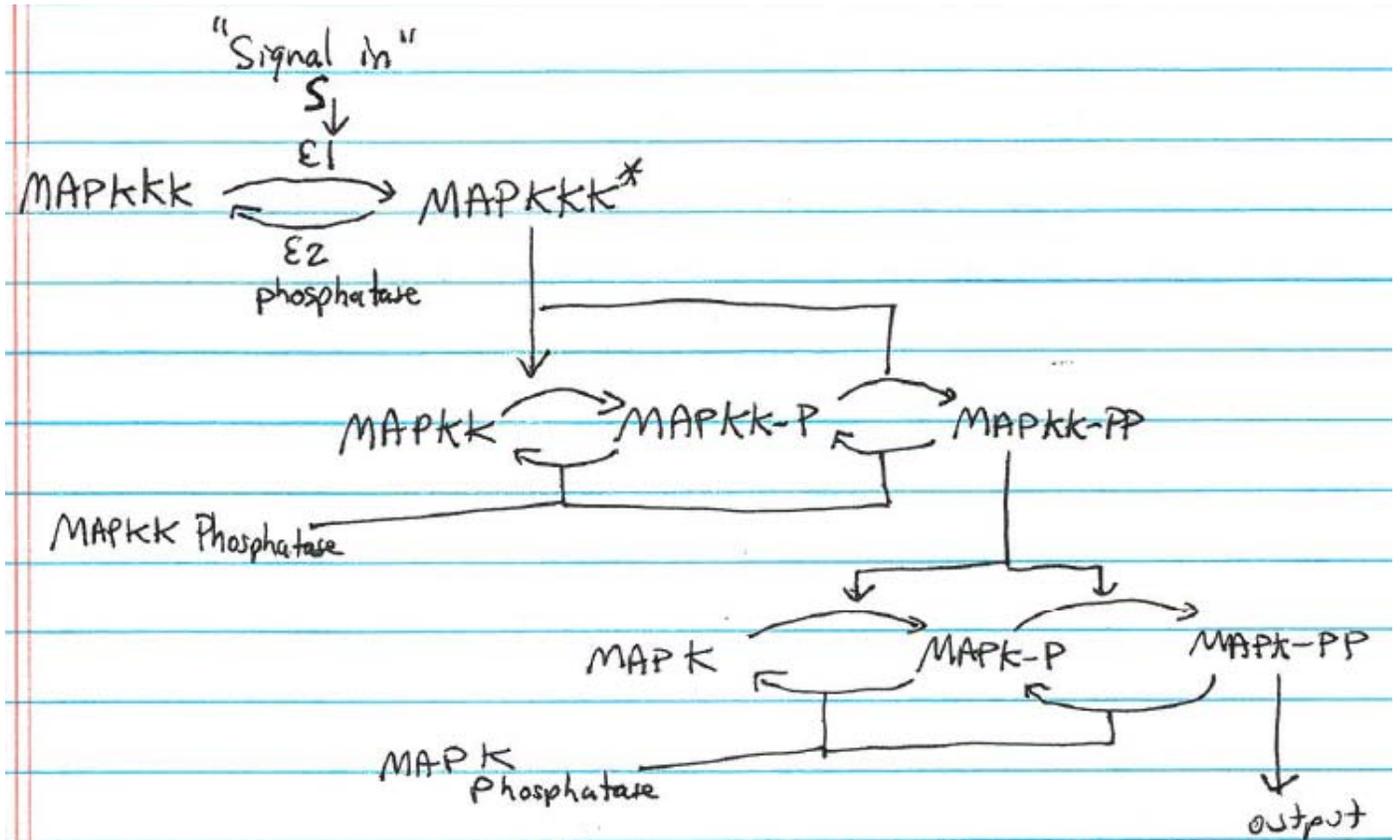
The first step in working with this diagram is to turn it into ODEs. The first step towards *that* is to write out the relevant chemical equations.

$$kkk^* + \varepsilon1 \underset{d_1}{\overset{a_1}{\rightleftharpoons}} kkk \cdot \varepsilon1 \overset{k_1}{\longrightarrow} kkk^* + \varepsilon1 \quad \text{activation of } kkk$$

$$kkk^* + \varepsilon2 \underset{d_2}{\overset{a_2}{\rightleftharpoons}} kkk^* \cdot \varepsilon2 \overset{k_2}{\longrightarrow} kkk + \varepsilon2 \quad \text{deactivation}$$

$$kkk^* + kk \underset{d_3}{\overset{a_3}{\rightleftharpoons}} kk \cdot kkk^* \overset{k_3}{\longrightarrow} kk\text{-}P + kkk^*$$

$$kk\text{-}P + kk\text{-}P'_{ase} \underset{d_4}{\overset{a_4}{\rightleftharpoons}} kk\text{-}P \cdot kk\text{-}P'_{ase} \overset{k_4}{\longrightarrow} kk + kk\text{-}P'_{ase}$$

$$\vdots$$
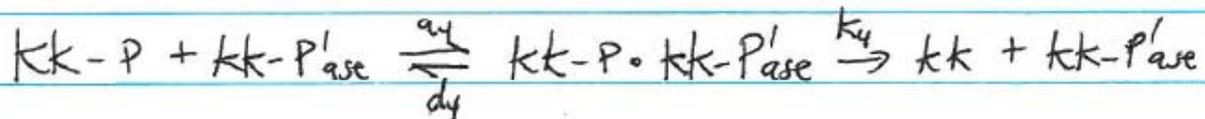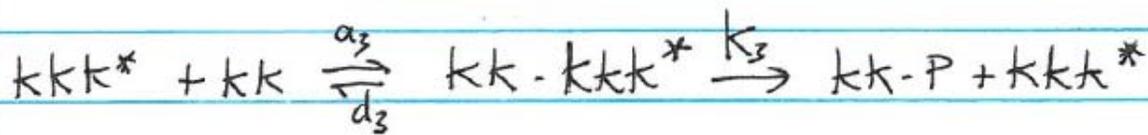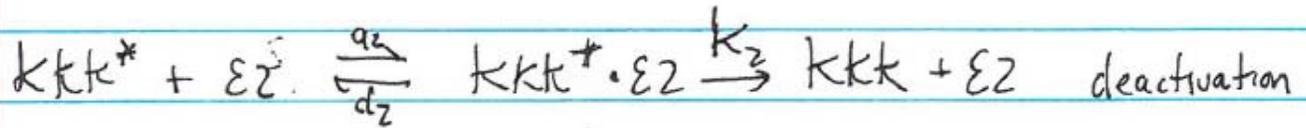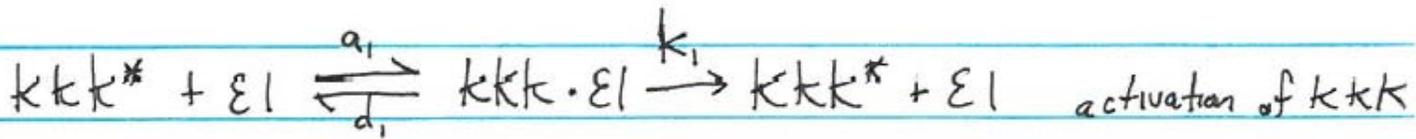
10 total reactions, all are listed out for your reading enjoyment in Huang & Ferrell, PNAS, 1996 (posted on-line)

A really good practice for your design project is to look at this model and try to re-create what Huang and Ferrel did when turning their diagram into chemical reactions and then into ODEs. See if you can come up with the same equations.

The ODES that result, then, are something like this:

1. $\dfrac{d[kkk]}{dt} = -a_1[kkk][\varepsilon1] + d_1[kkk \cdot \varepsilon1] + k_2[kkk^* \cdot \varepsilon2]$

2. $\dfrac{d[kkk \cdot \varepsilon1]}{dt} = a_1[kkk][E1] - (d_1 + k_1)[kkk \cdot \varepsilon_1]$

3. $\dfrac{d[kkk^*]}{dt} = -a_2[kkk^*][\varepsilon2] + d_2[kkk^* \cdot \varepsilon2] + k_1[kkk \cdot \varepsilon1] +$
$(k_2 + d_3)[kk \cdot kkk^*] - a_2[kk][kkk^*] +$

$(k_5 + d_5)[kk\text{-}P \cdot kkk^*] - a_5[kk\text{-}P][kkk^*]$

$\bullet \bullet \bullet$

18 total ODE's to describe this cascade.

And to this we add a handful of **Conservation Equations**, which keep track of how much enzyme we have in various states. They are the last equations we need in order to be able to solve for all the relevant variables.

$$[\varepsilon^*_{TOT}] = [\varepsilon 1] + [kkk \cdot \varepsilon 1]$$

$$[kkk_{TOT}] = [kkk] + [kkk^*] + [kkk \cdot \varepsilon 1] + [kkk^* \cdot \varepsilon 2] + [kkk^* \cdot kk]$$

$$+ [kkk^* \cdot kk\text{-}P]$$

Note, also, that there are several errors in the H&F paper. Beware…

Assumptions!

There are several built-in assumptions in this model. What are they?

1. **ATP bound to active kinases**
    a. We assume that it's always in excess. Adding it to our equations would monstrously increase our model complexity, and we can realistically assume that it's mostly in excess. Where can we not assume this? In the case where we bring in an ATP inhibitor, and ATP concentration is suddenly important. For the initial model, though, we can happily ignore it.

2. **Two-step distributive mechanism**
    a. This means that two phosphates are added onto the same substrate not at once, but after two different binding events.
    b. Why would the same kinase not phosphorylate the substrate twice in one go, rather than releasing and re-capturing the same substrate for a second phosphate addition?
    c. Kinases have a single reaction site in the substrate-binding groove, so it's not possible to add two different phosphates without letting go of the substrate. This has advantages.
    d. By forcing the system to phosphorylate two sites separately, we change the way that the cascade responds to the levels of external signal.
    e. This is caused by significant cooperativity, which allows the system to respond to minute changes in signal.

20.320 Analysis of Biomolecular and Cellular Systems
Fall 2012