

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

PROFESSOR: So today we're going to do an example of using some of the equilibrium concepts that we've learned to drug design. And it's going to be an example that's out in the literature. And potentially a very big deal if you can do drug design this way. And it's an example of how, remember how I told you that if you have the Gibbs free energy, you have everything and how clueless I was as a graduate student, and when I look back I think how silly it was for me not to realize how important it was. Well, this is an example where people calculate delta G's or Gibbs free energy changes for binding of proteins to receptors on cells, or ligands. And from that, design drugs that work much better than the real thing. And it's all about calculating delta G. We're going to see how that is, and how that comes in with equilibrium.

So the paper that this is based on was published in 2002. And since then there have been larger-scale trials, animal trials. And I believe there have been human trials of this concept, of this drug that they have designed. You've got the reference in the notes. And we're going to have to do a little bit of review of biology first, to figure out what's going on. And this particular process, this particular drug, is called the, I have it right here, granulocyte stimulating factor. Where was it in here? To give you the right name for it. There it is. It's a protein drug called GCSF. Granulocyte Colony Stimulating Factor. And it's there's wild tie for natural version of this protein. That is generated by normal tissue. And this protein goes to the blood, to the bone marrow. And it binds the receptors on cells that are in the blood and the bone marrow, and stimulates the growth of white blood cells, and stem cells, and also acts to tell the bone marrow to pump out these stem cells into the blood.

And the problem is that if you have a chemotherapy patient, their bone marrow cells are largely destroyed through the chemotherapy. And so they have a problem with the white blood cell counts, and with stem cell counts and all that stuff. So one of the ways that you can try to reverse this problem is by stimulating, artificially stimulating the growth or the proliferation of these white blood cells. And one of the ways you could do that is by giving the patient a lot of this protein, granulocyte colony stimulating factor, to stimulate the output of white blood cells from the bone marrow. Whatever is left of the bone marrow. To rebuild it up.

And so that's done with wild type, with protein. You can make the wild type protein. But if you had something that was basically the same thing but somehow mutated, where you made a few changes to the amino acid sequence, so that it worked a little bit better, then you might have a lot of people, a lot of patients. And that's what the basis of this paper is, that's referenced here. Is how to go about mutating this protein here to make it a drug that would

be more effective than the natural protein. So that you can give it to patients.

So now let me go back and tell you a little bit about this equilibrium that we're talking about. So we've been talking about equilibria of molecules coming together. And then creating new molecules, products and reactants. Well, when you're talking about ligands and receptors on cells, the same sort of ideas come in. The same equilibrium concepts come in. So if you have a cell membrane, the cell membrane could have a receptor in it. Which is a protein that extends through the membrane that's got some sort of pocket outside here. The membrane keeps on going. And you've got the blood on the outside here, or some of the tissue of the cells, and you've got the inside of the cell here. And this receptor is waiting for some ligand. A ligand being another protein that's floating around. That's specific for binding to this receptor here. And there's the ligand here. There's the receptor here. And every now and then one of these ligands will find a receptor and bind to it.

So we write, then, in equilibrium, ligand plus receptor goes to complex. This is happening all the time on cells. There are millions of different kinds of receptors, millions, billions of kinds of ligands. Wild types, fake ones, et cetera et cetera. When you've got a binding event, the cell knows that's something has bound, usually. And that usually triggers a cascade of other things. It signals a cell to do a lot of things. So, for instance, in the case of the GCSF, you've got the receptor on the cell. And you've got this granulocyte colony signaling factor, that would be the ligand here. It binds to the receptor on the cell. That triggers the bone marrow cells to produce more proteins, which signals more things to happen, and eventually down the way, after a bunch of signaling processes, out come a burst of stem cells or white blood cells. It's a complicated process, and every step of the way needs to be done correctly. And for this process at least, this protein here, this ligand protein, triggers the cascade. And so this binding and unbinding of this protein then triggers the whole sequence of reactions. So this equilibrium becomes super-important then.

OK. What else should we review about, I should have used a different board than this, but. This is going to get covered up. What else can we review about biology that would be interesting? So, other things that we will need to remember, need to know, is that these receptors on the surface of the cells aren't static. The cells recycle their receptors. Things can happen. Things get degraded. And so the cell is constantly taking these receptors, bringing them back inside the cell, merging them with lysosomes, where the pH is five, or 5.5 compared to the outside of the blood, which is 7.4. It breaks down the proteins into their amino acids, the cell can use the amino acids again to make more proteins. Make better, make new receptors. And that's how the cell recycles the receptors.

So basically, you end up with the cell taking the receptor. Making a small sort of cavity around the receptor. So there's the cell sitting here. There's the inside of the cell. There's the receptor being engulfed by the cell, now. And it could have the ligand in on it, also. There's the ligand sitting right here. It's called an endocytotic event, or it's endocytosis of the receptor. Into the cell. And then once you're inside the cell, let's put the nucleus in the middle

here, you've got things called lysosomes that are sitting nearby. There's a lysosome here. Lysosome. And there's your little vesicle that contains your receptor. Merges towards the lysosome, these two get together.

Then you have your ligand, I'm running out of colors here. So there's the receptor and the ligand together in here. And then the pH here becomes 5.5, things break apart. And you've got to find another receptor, another ligand, to continue the process. Everybody knows this biology, or you know it well enough. OK, good.

So let's go do this example. So this is the process of binding. We can have a  $\Delta G$  associated with this,  $\Delta G_a$ ,  $\Delta G_0$  is  $-\text{RT} \log K_{\text{sub } a}$ . This is the association equilibrium. And you can have the reverse process, where the complex gets broken up into receptor plus a ligand. And then you have a  $\Delta G$  for this process here. So this is  $\Delta G_d$ , let's call it. This would be  $\Delta G_D$ , which is the negative  $-\text{RT} \log K_{\text{sub } D}$ . This is the dissociation process. And the dissociation process,  $K_{\text{sub } D}$ , is equal to  $R$ , the concentration of the receptors times the concentration of the ligand, divided by the concentration of the complex. And the lower  $K_{\text{sub } D}$  is, the smaller this number is, small means that you're mostly on this side here, mostly in the complex, the tighter the binding. So small  $K_D$  means tight binding.

So if I want, in principle, then, if I want to design a drug that's going to signal this event here, I want  $K_{\text{sub } D}$  for this ligand that I'm going to design to be very small. To be small. Small enough so that it binds strongly and does its job, so I don't need too much of it. Now, you need to do experiments to figure out what's going on. So how do you do these experiments? You need to be able to measure, then these  $K_{\text{sub } D}$ 's, experimentally, to see whether or not what you've designed on the computer, when you calculate  $\Delta G$ 's on the computer you've got to know whether or not it's working. And this is still an experimental science.

How does it work? So, usually you do the experiment with the ligand concentration very high. Very large. So that the concentration of the ligand at any time is basically the same thing as the concentration of the ligand you put in. So you overwhelm the system with ligands. So that  $L$  is much bigger than  $C$  or  $R$ , and so it doesn't matter which way the equilibrium is.  $L$  stays roughly the same.

So you know, throughout the experiment, what concentration is. Then you get a bunch of cells. In a well or something, or a 96 well plate with a bunch of different kinds of ligands. And you can take your ligand, you can label it radioactively. So you can take your ligand as some long protein. And you can take an iodine 125, let's say, radioactive label on your ligand, on your protein. So you can keep track of it. The nice thing about radioactive labels, and which is why people use them in biology or bio-medicine. We use them to look at, for instance, bio-distribution of things in animals. We want to know where things are, and whether they all left the animal, or. Because you can use a Geiger counter and you can count the events. The radioactive events. And that gives you extremely quantitative analysis of where things are. So you take those radioactive ligand. And,  $L_0$ , very large

concentration. You expose this concentration to the cells. The cells have some receptors on the surface. There's a bunch of receptors on the surface of the cell. And some fraction of the receptors will have the ligand bound to them.

So you incubate. You let it wait a while. Then you wash. So you get rid of all the excess ligand that's on top there. And then you take your petri dish, or your 96 well plate. And you read the radioactivity that's coming from the cells. And that signal, that radioactive signals, tells you exactly how many ligand you have on these cells. You knew what the cell concentration was initially. So that tells you what the concentration of ligands, of complexes, was.

So this experiment then gets you, experimentally, gets you the concentration of C, the complex. So now this is something you know. You also know what L<sub>0</sub> was, because that's what you put in. And that's enough for you to find out what K<sub>D</sub> is. And if you know what K<sub>D</sub> is, then you know what  $\Delta G^0$  is. And so generally, then, let's go ahead and do that.

So we know what L is. We know what L<sub>0</sub>, C is, let me just do it on this board here. So we start with rewriting K<sub>D</sub> as equal to R times L divided by the complex concentration. And now L is basically L<sub>0</sub>, so we can use that approximation. So we have L<sub>0</sub> sitting here. We still have the complex on the bottom, and that's something that we've experimentally discovered. And the concentration of receptors, this is the concentration of receptors that don't have anything bound to them. So it's these guys right here. That's equal to the concentration of receptors, the total number of receptors, minus those receptors that have a ligand bound to them. Complexes. Something we can experimentally define, or discover.

So we replace this R here with R<sub>T</sub> minus C. And then we're all at equilibrium. So we're going to put a little equilibrium sign under these C's here. equilibrium. And then we can rearrange this equation so that, people like to have plots that are linear, in a way that is a linear plot, where the slope of the plot is the inverse of the equilibrium constant.

So you do some massaging of this equation here. And you rewrite it in terms of C over L<sub>0</sub>. There's the equilibrium concentration of the complex. This is the total receptor concentration divided by the equilibrium constant. And then you have the equilibrium concentration of the complex divided by K<sub>D</sub>. And so you plot, then, this ratio. Which is an experimental ratio. You've just measured this concentration of complexes using this radioactivity, radioactive labeling experiment. You know what this is, because that's what you've put in. This is your x-axis on your plot. This is what you've measured. And this is going to be the slope. It's called a Scatchard plot. After Mr. Scatchard

So you get a straight line. So we're plotting here the equilibrium constant of the complex. And on this here we're plotting the ratio of the complex divided by L<sub>0</sub>. And we get this straight line like this, where the slope is one over K<sub>D</sub>. Minus one over K<sub>D</sub>.

OK, now there's a couple of things that you can look at that are sometimes useful. In this analysis here. You can also rewrite this equation up here in terms of the ratio of C equilibrium divided by RT. So that's basically the ratio of receptors that have a ligand attached to them, divided by the total concentration of receptors. So if most of the receptors are empty, then this is a small number. Then most of the receptors are taken up, this is a number close to one. It can't be anywhere, it can't be ever bigger than one, because the biggest number of complexes you can get is the total number of receptors. So if you take this equation here and you massage it a little bit, one over one plus KD over L0, and that, you can also plot that. As a function of L0. How much ligands you put in. And you find this is something that saturates at one.

So this ratio here is going to saturate at one. This is C equilibrium divided by total number of receptors. And so if L0 is small enough, if L0 is small enough, meaning that it's smaller than KD, so if L0 is much smaller than KD, then you can rearrange this ratio here so that C equilibrium divided by RT is roughly L0 over KD. So that's linear, with a slope of one over KD. So this slope here is one over KD is the slope.

And as you get L0 to be quite large, bigger than KD, then this saturate to one. And this one over something very large become zero, and basically you end up with something close to one. So you saturate at one. So that's another way of doing this. But this is really what we want to concentrate here. The fact that you can get KD out of experimentally. If you have KD, you have delta G.

So now let's go back and think about this whole process here. If we're going to design this drug here, this protein, artificially. So what you want, then, is you want something that's going to bind strongly enough to receptor, to stimulate the growth of granulocytes, or the colony of granulocytes. But when the cell decides to recycle the receptor, and destroy it, chew it apart in the lysosome here, you want this drug not to be degraded. Because you have to keep injecting in the patient. So you want this drug to release from the receptor, before the lysosome has a chance to chew it up. So that means that you want the equilibrium constant at pH 5.5, you want that KD, to be much larger at 5.5 than you want it at 7.4. You want strong binding at pH 7.4, and you want weak binding at pH 5.5. That way the drug gets recycled. Can find, go back to the blood. Bind to a receptor again, generate the signaling events, gets engulfed by the cell. Releases before it has a chance to be chewed up, et cetera. The wild type will get chewed up. The regular kind will get chewed up, and so that's why you have to, with the patients you have to keep giving them this drug over and over again. Because it gets chewed up by the cells.

So that's the design principle that these authors had in mind when they started their study. And so they knew what the sequence, what the amino acid sequence was, for this wild type drug, and they started doing point mutations. Changing one amino acid here and there. And cranking out the calculation to calculate delta G for binding of this protein to this receptor. At pH 7.4 and at pH 5.5, and getting differences of delta G's.

So let me then review again. What is the motivation here. The motivation is to try to get the ratio,  $K_D$ , at 5.5 divided by  $K_D$  at 7.4, pH 7.4. And you want this to be bigger than one. And really as large as possible. As possible, of course within limits. So we want this to bind at pH 7.4. And the point of comparison is the wild types. Which at 7.4 has a  $K_D$  of 270, roughly. And at the ratio of pH 5.5 to 7.4 is 1.7. So it's a little bit weaker binding at 5.5. Remember, large  $K_D$  means weak binding. Small  $K_D$  means tight binding. So the fact that  $K_D$  at 5.5 is bigger than  $K_D$  at 7.5 means that it's slightly weaker binding at 5.5 than 7.4.

So this is sort of the baseline that the protein designers had to deal with. So everything gets compared to this guy here.

So they don't actually, in a calculation they don't actually measure this. What they do is, they look at delta G's. And so they look at differences of delta G's. They look at delta  $G_0$  for the binding of the ligand to the receptor. At 7.4 minus the delta  $G_0$  at pH 5.5, let's call this the delta delta G. Since you know delta G is minus  $RT \log K$ , this is equal to minus the delta of log  $K_D$ , which is log  $K_D$  at 5.5 divided by  $K_D$  at 7.4. So then they calculate delta delta  $G_0$ , and it's completely related to this ratio that you measured in the experiment. And so for the wild type, this delta delta G is just basically the log of this number here. Is 0.53. So now they go ahead and do their calculation. And they find a couple of mutants. Where, and they focus on delta delta  $G_0$ . They want something that's bigger than this. And then they'll worry about whether there are more problems associated with it. So they found two mutants, let's call them D110H and D113H. I think it's a histidine mutation, point mutation at 110 and 113, where this ratio here was 8.3 and 17. So, huge differences here. More than a factor of 20, or factor of 30 difference in the change in the binding efficiency, at least according to these delta delta G, between pH 5.5 and pH 7.4.

That means that this protein here with the point mutation, with this one point mutation, if it binds as strongly to the receptor on the surface, as soon as the lysosome comes in and begins to decrease pH inside the cell, this ligand is going to come off. And it's going to be able to float away. Hopefully, before it gets recycled. Before it gets chewed up. So they can be used again. So that's good.

Alright, so experiments were done. And on the D110H and the D113H. And  $K_D$  was measured in the experiments. And remember, the wild type is 270. The  $K_D$ 's were at 370. And 320, with some error bar. And the different, the ratios of  $K_D$ 's were measured. There were 4.4 and 6.8. With some error bars. And they turned out to match reasonably well, at least as far as comparison between experiment and theory. Still not great for these sorts of calculations, because pretty involved. A lot of approximations go on in there. And it gives you a rough guide. So this 17 here is probably not quite right, because it doesn't quite translate to 6.8 here. This difference between 4.4 and 6.8 doesn't quite match the difference that they saw here in the calculation. But it's the right direction, right? And that's why you still need to do experiments.

This number here is a little bit bigger than this number here. Which means that these don't bind quite as strongly. Because  $K_D$  small means strong binding. That means the initial binding of the ligand to the receptor is not quite as good as the wild type. That's often the case. It's often not so easy to find something that binds as strongly as the wild version. But, it's good enough. And this ratio is really what clamps the deal in this case here.

And so, these mutants, in fact these two mutants are the ones that have undergone the animal trials. That are in the pipeline. This is a big deal. This is a big deal because this is a huge, huge market. There's a very large number of people that are affected with chemotherapy. And this is basic thermodynamics here. It's basic thermodynamic calculation of a complicated molecule with some fairly simple equilibrium concepts. OK, any questions on this?

OK, well, we're ended really early today. I don't have the phase transition ready, but Keith Nelson is going to start a phase transition next time.