JOANNE STUBBE: So anyhow, this is where we ended the lecture last time. We're finishing up. We were talking about two kinds of regulation for cholesterol sensing. One using the transcription factor sterol responsive element binding protein, and how we can use that when cholesterol levels-- or sterol levels are low to upregulate the amount of sterol by turning on the genes for the biosynthetic pathway. And so that's what we were talking about at the end, and this week in recitation as well. And also you can turn on the gene for the LDL receptor, which then allows you to take more cholesterol in from the diet.

And so what I wanted to do is in this week's problem set 7, you were focused on looking at the SCAP protein and how do you know what the SCAP protein is doing. And what you were seeing in the data you were give given, which was taken from one paper. There are many papers trying to study these proteins to understand what's going on with these complex membrane proteins.

Where does the sterol bind? How does the sterol bind? What is it that causes this chemistry to happen where this complex migrates from the endoplasmic reticulum to the Golgi, allows cleavage chemistry to happen, and ultimately a little piece of DNA which binds to the sterol responsive element to actually turn on a bunch of genes that we just talked about?

So I just wanted to say one or two words about the players. And you've all thought about the players by now. We're going to come back and look at some of them in a few minutes, but the key player in the scheme I just showed you was SCAP and that was the focus of what you guys had to do in your problem set. And what you notice again is a sterol sensing domain. And there is also-- I point out we'll come back to that at the end of today's lecture-- there's a sterol sensing domain in HMG-CoA reductase, which we will see is involved in post-transcriptional regulation. So we're still looking at transcriptional regulation now.

And the question is-- and then you have a bunch of other transhelices, single transmembrane helices probably helical within the membrane. And the question is, how does this guy work to

allow the model we showed on the previous slide? So that defines all those terms. Hopefully, you're now familiar with all those terms. It's written down in a place we can go read about it again. But this is the cartoon that you were dealing with in the problem set.

And so the key question is-- most of these things are defined. Whether you have a transmembrane helix is defined by some kind of sequence gazing, and then you have to do experiments to test whether the model is correct. And we don't have any pictures of the SCAP protein at all at this stage. So the kinds of experiments that you were looking at your problem set are the kinds of experiments that people are still doing to try to figure out how all this information is coordinated to allow the chemistry to happen, or that migration from the ER to the Golgi.

And we talked about last time-- we talked about a zip code. So if we look at SCAP, so we have eight transmembrane helices. And the key to the way this works is that there's a little zip code. And you've seen a zip code before transiently when we were looking at a zip code on the LDL receptor, which targeted AP2 to then bring in the clathrin coats to make the clathrin-coated vesicles.

I think what you'll see over the course of the rest of the semester, there are lots of times three or four amino acid sequences that are the key that allow some kind of confirmational change to occur, which can trigger off a sequence of events that people have found by doing a lot of studies on the system. So the zip code here, and that's what you were focused on in your problem set, again. And I don't expect you to remember any of this, except to sort of know that these little zip codes play a role quite frequently in biological transformations of these complex systems. And so here is the little zip code.

And we've been talking about so far what happens under low sterol levels, where we want to make cholesterol or get more in from the diet. And under these conditions, if we want to make more cholesterol, we have to turn on the biosynthetic machinery, HMG-CoA reductase, or turn on LDL receptor that allows you to take things into the cell. And so this is proposed to be a key player in loop 6.

And this loop 6, which is pretty big-- and you might ask yourself the question, where did they get this loop? So that's something that you have to design experiments to figure out what is cytoplasmic, what faces the lumen, how big are these loops? All of that plays a role in thinking about how this works from a molecular perspective. So loop 6 plays a key role, as does, you

can see from your problem set, loop 1.

So what does loop 6 do in the model? The cartoon is shown over here. Under these conditions where you have low sterol, is below whatever the membrane concentration is that you looked at for the recitation this week, we're down 3% or less or something like that, what happens is this little zip code is exposed. And it in some way recruits proteins that are involved in another complex process that we aren't going to talk about where you can bud off little vesicles where proteins of interest and also lipids can be moved from one membrane to another, the ER in this case to the Golgi. And so this interaction is a GTPase here.

There are a couple of additional proteins that have been identified. We're not going to talk about the details. But this is the key to allow movement into the Golgi, which then you have the defined proteases that we've talked about before which allow cutting and allow the little piece with the helix loop helix at the N-terminus to become soluble and then move to the cytosol.

So we want to ask the question now and spend a little bit of time, what happens with high sterol concentrations? So everything we've looked at has been-- this is at low sterol concentrations. And under these conditions, this zip code targets SCAP and SRE-BP to the Golgi.

So now we want to go to the second set of conditions. And again, in this week's recitation, we're focused on high sterol, low sterol concentrations. High sterol, what is it that allows this to prevent movement into the Golgi so you can get this processing so you can turn on HMG reductase and LDL receptor biosynthesis. And so the proposal has been-- so in some way this is connected to sterol, so we're at 5% or 6% sterol in the ER membrane.

That's what we're going to discuss today, for those of you who haven't had recitation. How do you know what the turn on versus the turn off is for sterol levels? Where does it bind? How does it bind? All of that becomes pretty interesting. And in high sterol we now introduce yet another player. So in addition to SCAP we now have to pay attention to INSIG. So that's the other player. We'll see that INSIG is this protein that's the linchpin for all the regulatory mechanisms.

So if we go back here, what do we see about INSIG? It's small. It's much smaller than SCAP. SCAP is this huge protein, 1,200 amino acids. Here's 200 amino acids, all transmembrane. Recently, actually, there was a structure not of the human system, but of a bacterial system. It's not found in the vast majority of bacteria, but they found one and so they've gotten some--

they have proposed some model for how INSIG could be in the membrane.

Now, one of the things that I find confusing to think about the molecular basis for what's going on-- which we don't know anything. You're looking at a cartoon level-- is we have so many transmembrane helices, but is this thing a monomer? Is it a dimer? Is it a tetramer? Is it a hexamer? And how do you look at that? Because when you solubilize it, you have to put it in detergent, et cetera. And is what you see in the test tube, how do you relate it back to what's going on inside the cell? And I think we really don't know.

Most of these things-- both SCAP, which is thought to be a tetramer, and now in this new paper they're claiming it's a trimer of dimers. Just add to the complexity of trying to figure out how all these things interact. So that's the issue with doing these kinds of experiments. We don't have very good experiments. We need people to invent new ways of trying to ferret out how these things interact within the membrane.

So here INSIG is going to be a key player and SCAP is also a key player. And so somehow, in the presence of sterol, so we're at high sterol, in the presence of INSIG, we need to get rid of the zip code. That's the bottom line. So the model is in the presence of sterol, we remove-- we don't really remove it, but we hide the zip code.

And when you hide the zip code, so this is shown-- I'm not going to draw this out on the board because I think they draw out better than I can do it, and we really don't know what that's going on anyhow. But again, this comes from a region where it's accessible to another region where these proteins can no longer bind to do the transport of these proteins into the Golgi. So that means you never get processing of SRE-BP to become active. Does everybody understand that?

So one of the questions is, if I asked you to design an experiment, hopefully you're now starting to be able to think about designing experiments. Is there any kind of an experiment you might think you could do to-- a simple experiment that you might be able to do or try to do that would allow you to show that you you'd undergone a conformational change in a loop 6. Anybody think of something?

So loop 6 this is big, huge-- it's proposed to be a big, huge piece of polypeptide and it's proposed to undergo-- this is a cartoon, but a tremendous conformational change. And so what you need is some kind of a simple probe that might tell you that it's undergone a conformational change. And what might some kind of a probe like that be? How could you

design something like that? Yeah?

AUDIENCE: A FRET experiment?

JOANNE STUBBE: OK, so that would be one thing you could do. You couldn't just do a FRET experiment to do it. What would you have to do?

AUDIENCE: Incorporate fluorophores.

JOANNE STUBBE: Yeah, so you'd have to incorporate fluorophores. So the issue with FRET is not only now do you have one problem, so you can put it one place, but now you've got to figure out where to put the second FRET. So that would be a lot of experiments, but now you can do mutagenesis, so you could probably do an experiment like that.

Is there anything else you can think about? Yeah?

AUDIENCE: So I don't know how much is known about the conformational rigidity of loop 6 and also I'm not sure-- I've ever seen this in a membrane protein, but you can maybe look at proton exchange, like [INAUDIBLE] backbone exchange.

JOANNE STUBBE: So that's a sophisticated experiment and I would say there probably wouldn't be my first choice. But the idea that it's accessible, is there any kind of an enzyme that you might be able to use that could-- again, you'd have to be lucky, but you could look at the sequence and think about this. Is there any kind of an enzyme you might be able to use that could sense a change in the conformation? And if the model is right, it's on the cytoplasmic face.

And so the answer is people use trypsin. So if you go back and look at the sequence, there's an arginine. And actually, I wouldn't have expected you to see this, although I think it is mentioned in one of the papers where it becomes much more accessible in one state than the other, so you get proteolytic clipping. But those are the kinds of things-- what other kind of experiment could you do? So you think you're undergoing a conformational change, what kinds of probes did Liz talk about that might allow you to see some kind of change in conformation? So we have fluorescence probes, which we haven't really talked about yet. Can you think of what other kinds of probes? She spent a whole recitation on it.

AUDIENCE: [INAUDIBLE].

JOANNE STUBBE: The what?

AUDIENCE: As far as binding [INAUDIBLE], you would do it in cross-linking.

JOANNE STUBBE: Yeah, so some kind of cross-linking. You might get information out of that. To do that again, you've got to put in the cysteine. So here you might have issues because you have all these cysteines down there. So the question is, could you do that? Cysteine is the most easy-- that's the easiest side chain to modify. On the other hand, these cysteine's really playing a functional role, there's no way you could modify it. But those are the kinds of things. Hopefully you're starting to get a battery of tools that you're learning about with different systems that you could further probe this, but you can just see that this is, I think, a really challenging problem.

So the idea is it does undergo a conformational change and you can no longer see the zip code. It disappears. And it resides, and so the model, then, becomes here that the sterol binds to the sterol sensitive binding domain and recruits INSIG. Is that true? Could INSIG bind first and that the two of them together? And in the paper that you were looking at in recitation this week, the concentrations of INSIG were elevated and they got a result that you may or may not have been able to predict what the outcome would be.

So again, the order of binding and interaction, I think, really still remains something that needs to be studied. So we just form a complex in this case. So we form a complex. So in the membrane-- we have our ER membrane. We have INSIG. I'm not going to draw the-- we have INSIG. It overlaps somehow with SCAP.

And then SCAP somehow binds the sterol, so you need to have a sterol here. And this is a sterol-sensitive domain. And then up here you have a loop and the zip code hides. So that's the model. And so that's as sophisticated as we can get at this stage, which is, for a chemist, not particularly sophisticated. But that's still the model. The model has been in the literature for a long time and we don't know that much about it.

Here is just another cartoon from another paper where, again, the little gold thing here is the INSIG. This is the SCAP. Here is the sterol responsive binding element. And here they don't show the hiding of the little zip code. So this is the model that people have put forward for how low cholesterol levels allow you to turn on the genes required to make more cholesterol or to take it from the diet. So that's a transcription model.

So in the last lecture now what I want to focus on is post-transcriptional regulation. So this is lecture 5. We're talking about post-transcriptional. And what you're going to see, we're going

to focus on again is our common player in both of these regulatory mechanisms. INSIG played a key role in keeping SRE-BP in the ER membrane. We're going to say INSIG plays a key role also in destroying-- so at high cholesterol concentrations, you don't want to make any more HMG-CoA reductase. And if you have a lot of it in the membrane, you want to get rid of it. So INSIG and HMG-R are going to interact with each other based on-- so this is a high sterol.

And what's going to happen is, in a sophisticated way, HMG-R, which makes cholesterol, is targeted for degradation. So this is the model. We're going to come back to this model over and over again. Sorry. My handwriting is getting worse and worse.

So we have INSIG. We have HMG-R. We want to get rid of it because we don't want to make anymore cholesterol. That's the bottom line. So INSIG is a player in both of these mechanisms.

So what I want to do now in lecture 5-- that's what this is here. What I want to do in lecture 5 is really talk about how you do degradation in eukaryotic cells in general. And then what I'm going to do is come back and ask the question, how is HMG-R, HMG-CoA reductase targeted for degradation.

So that's the overview of where we're going. It's pretty simple and we'll see that mechanism of the proteosome, it's much more complicated, but there are many similarities between this chamber of doom and the one you saw with ClpX and ClpP where you've spent a lot of time discussing what we know about it from a more chemical and biochemical perspective.

And what I'm going to show you at the very end is not only does HMG-CoA reductase get targeted, but now in the last couple of years, they found that all of the proteins involved in cholesterol homeostasis get targeted by different mechanisms to get degraded. So protein degradation inside the cell is extremely complicated. So what I'm going to do is give you an outline of a generic picture of how it gets degraded. And the caveat is this is a very active area of research and I think you'll see why it's so complicated.

And then the next problem set, problem set 8, there's been a model in the literature that this is the equipment that targets this protein for degradation. And I'm going to give you a bunch of data that says that may not be correct. So this is what you're dealing with. Every time you pick up a journal, there's another model and perhaps there are five or six different ways-- not that many-- three or four different ways that you can mediate the degradation. And we're in the process of trying to unravel this.

So this is where we're going. And so what I want to do is start out by looking at-- and then if we have time at the end, I'll come back to both recitations. But I probably won't have time at the end because I want to move on to the next modules. But you'll see that the recitations really are pretty much linked to what we're talking about in class. So it's unfortunate that they weren't timed a little better, but that's the way life is when you're trying to balance all of these acts.

So this is the overview. I'm not going to draw this on the board, but I'm going to walk through it step by step. So this is a cartoon overview. You can see it's pretty old. We've learned some stuff, but there's a lot of stuff that remains unknown. So let's just work through the cartoon and then we'll walk through who the players are, what the model is and then in the end, we're going to return to HMG-CoA reductase.

So we have a protein and we need to target it for degradation. How does anything know that this is targeted for degradation? The protein's the same in the beginning. In the end, how do we know why this protein has a different kind of half life than some other protein?

We haven't discussed that, but we asked the same question in bacterial systems and I'm going to spend not very much time on it. But the N-terminus of the protein can be modified in many ways. This is called the N-end rule, totally mind boggling. I might give you a few examples of this on a problem set. But you can add on amino acids or take off amino acids. It changes the lifetime of the protein from minutes to hours.

So this is like-- when this first came out, I, said there's no way that can be true. So we're talking about a few amino acids, just like we're talking about these zip codes over here. It's true. And the way the rules work or have evolved, they're different in all organisms. They're in all organisms, but they're all distinct.

So another way that I think is key, and we're still trying to figure this out is that many proteins, are post translationally modified by phosphorylation or hydroxylation or whatever. I think that's also a key thing that's going to target them for degradation. So we have a protein. Somehow it's going to get targeted for degradation. What does that?

So it turns out we're going to be introduced to a molecule you saw in your first recitation, ubiquitin, small little protein, like a rock, 76 amino acids. What is it doing? It's like the SSRA tag except more complicated, that you saw before. And then we're going to be introduced to three proteins-- E1, E2, and E3, an activating enzyme, a conjugating enzyme, and a ligase. And I'll

sort of define for you what the function of these proteins are. And you'll see that they require energy.

Maybe not surprising, the Nobel Prize was given for the work on discovery of how this little system works a number of years ago, where a major player in that was a mechanistic entomologist named Ernie Rose, who nobody ever heard of. I remember when the Nobel Prize came out, the chemist was saying, who the hell is this guy? Well, so that's because, again, they don't care about how enzymes work.

But what's amazing is this guy is one of the most brilliant people. He's dead now, but he's one of the most brilliant people I've ever met. And he was [INAUDIBLE]-- he did thousands of things that were really creative and important, but this is the one, because he was hooked in with the guys that were doing the biology, that allowed him to elucidate what was going on.

So we're going to take this a little ubiquitin and somehow this equipment is going to attach the ubiquitin onto the protein that's targeted for degradation. And we'll see that you have to have multiple ubiquitins attached to get degraded. That being said, we now know that almost all proteins can be ubiquitinated. We know ubiquitin has something like 20 homologues, look alikes, and they all do different biology. So this is another example of post-translational modification. We're only going to focus on targeting for degradation. That's what I'm going to show you but the ubiquitinome is quite complicated.

So once it gets the ubiquitins attached, what do you see here? You see the proteosome. This is the chamber of doom. We'll come back and look at that, just like clip x and clip p. So you have clip p here, the chamber of doom. And then you have little pieces on the top and the bottom of that, which would be sort of like clip x. Hexameric ATPase we'll see is much more complicated in human cells.

And so what do you have to do? You have to unfold the protein. You have to thread it into the chamber of doom. You have to break it down into pieces and you spit out the pieces. This process requires ATP like you studied in the bacterial proteosome. And then there are actually many different proteosomes in human cells and I'm just going to talk about the generic proteosome. So that's the cartoon overview.

So I want to say a few things about-- so let's start by looking at the proteosome. And again, this is the human proteosome. And if you look at these big machines, you've already learned one way you characterize them is by their sedimentation in some kind of a centrifugal field.

And so these things migrate. Like, a 26S particle only has a sedimentation value of 26. So it's huge, and it's 2.5 megadaltons. So this is a huge machine just like the ones you've been studying in the first part of the course.

So it turns out this can be divided into two parts as you've already seen and you can see over there. You have the 20S, which is the core proteosome. And then you have a 19S lid. Actually, you can have multiple lids and in these lids there can be 20 proteins, 15 to 20 proteins. So the lid contain 15 to 20 proteins. We'll come back and look at this a little bit. And so this is going to be in-- and among these things are the AAA plus ATPases, which are actually quite distinct from what you're going to see, what you have seen in the bacterial proteosome.

So here again is going to be the 20S core. Here are the proteins, so this 20S core. Here are the proteins involved in the lid. Some are tightly bound, some are not tightly bound. Remember, we had a hexameric ATPase, so RP-- I can't remember the acronyms-- RPT, and there are six different ATPases, not one, six. But they form a hexameric structure.

And then you have a lot of additional proteins that we're going to come back and look at, but one of you might expect would be something that could recognize ubiquitin, just like you had something that recognized the SSRA tag. It turns out ubiquitin is recycled inside the cell. So the equipment that allows you to cut off the ubiquitin so it can be used again, de-ubiquinating enzymes, is also located in the lid. And you can also imagine that could be many kinds of adapter proteins because we're going to be able to degrade many, many, many proteins under different sets of conditions. So this changes in composition, as opposed to the chamber of doom, the 20S proteasome.

So let's look again at the core particle, so the core protease. Let's abbreviate it CP. And what do we know about this? What we know is the following-- that it forms four heptameric rings. And the rings, so each one of these is a 7-mer.

And it turns out we have two kinds. They're actually pretty similar to each other, just like the proteosome from bacteria. But we have alpha, we have beta, we have beta, and we have alpha. And we call them-- we put i's next to them because, again, they're not the same. So they're all different. So they're all structurally the same, but they're all different. So i can be 1 through 7.

So what do we know from studies that people have done? The key thing is alpha. So these alphas at the top and the bottom are inactive in terms of chemistry of peptide bond hydrolysis.

So all of the chemistry-- so these are each the beta heptamer is also-- these are in the center. These are active. So the activity is here and it's flanked by to heptameric rings that are inactive.

And so what do we know about beta? So even though we have beta i, where this is 1 through 7, it turns out that four out of the seven betas are inactive. So again, you saw the complexity with Saunders' talk on single molecule stuff on ClpP, right? So four of the seven betas are inactive. So that might not be so different. But I think every proteosome, even though the architecture is sort of similar, has evolved slightly different strategies to deal with the same problem.

But what's interesting here is, it doesn't matter which one is which, but the three betas that are active all have different specificities of peptide bond hydrolysis. So B1 has D,E specificity. Hopefully you all know what that means. That means simply, for example, if you had an aspartate and this is where the peptide bond cleavages, they recognize aspartate in the P1 binding site.

So if they recognize aspartate and a glutamate, B2-- or I might have the numbers mixed up-- recognize lysine and arginine. What does that look like? We've seen this now a hundred times. That should remind you of trypsin. So we have yet another lysine-arginine-dependent protease, and these are all over the place in the body. So it's not just this one little site. It is, maybe, in the proteosome. But if you look at blood coagulation, there was something like 15 lysine-dependent proteases, and they've got to all be controlled, otherwise we would clot all the time. Yeah?

AUDIENCE: When you say four of the seven [INAUDIBLE] are inactive.

JOANNE STUBBE: Yeah.

AUDIENCE: Do you mean that--

JOANNE STUBBE: They can't catalyze any peptide bond.

AUDIENCE: But is it, like, in some of the other proteases we saw where it changes? Or is it for a given or a specific molecule, a specific protease it's always the same four units that are inactive?

JOANNE STUBBE: It's always the same four that are inactive, but whether they're locate-- how do you call 1, 2, 3, 4, and how they assemble? An interesting question that actually people are studying in

thermophilic bacteria. But you can imagine if they had be-- I don't think they have to be predisposed. That's why I'm saying the numbers don't make that much difference.

AUDIENCE: [INAUDIBLE].

JOANNE STUBBE: So it doesn't have to be B1, B2, B3, B4, B5, and you always see the same. I don't think that's true, but I don't really know. So you have a different specificity there.

And the third one, which they call-- I don't remember what they call it. They call it B4 in the paper, so maybe they more know more about this than I do, but you have hydrophobics and you have aromatics. Where have you seen that kind of a protease before? Yeah, kind of with trypsin.

So you're seeing this-- these are the common proteases you find all over the place. I mean, we use them as tools all the time as biochemists, these three. There are many variations on this theme. So anyhow, so what you have then is basically heptameric units where the activity is here. These are inactive. And somehow you have to get the protein that's going to be degraded just like you did in the clip p protein, get it into the chamber of doom.

So what do we know about the mechanism of cleavage? So I'm not going to go through the mechanism in detail, but the mechanism I'm going to say a few things. The mechanism of peptide bond cleavage is distinct in that what did you see in clip p? You had a serine-type protease that involved covalent catalysis.

Here, what you have in the human system is a threonine. So that's sort of unusual. There were been a number of these since this was discovered a while back. There have been a number of threonine proteases, so this is the rest of the proteosome.

So it turns out that threonine is at the N-terminus of the proteosome. So that becomes important in terms of its chemistry. So this is the N-terminus. And the two things that you could picture that might be involved in catalysis, based on what you've learned about the bacterial protease, is that you have a serine, you have a threonine. They both have OHs that could be involved in covalent catalysis. So this OH is thought to be involved in covalent catalysis.

And remember, what do you have and what do you have in the case of the clip p protein? What also is required besides a serine? You need some kind of general acid, some kind of general base catalyst in all these proteases. In the case of serine proteases, it's usually

histadine. It's not a histadine in this case. It is the N-terminal amino group of the protein that's proposed to be the N-terminal amino group of the protein.

Now, if you look through the PKAs amino groups at the N-terminus versus lysine, they're always lower. So this amino group, so the N-terminus amino group is thought to be the general base catalyst and the general acid catalyst in the mechanism. So that's the proposal.

And I must say, I don't think we know a whole heck of a lot. I haven't read the literature [INAUDIBLE]. There's not that many people working on the mechanism at this stage. So this is the proposed mechanism, so just put it in quotes, "proposed". And it's completely sort of analogous to what you went through in the first part of the first part of the semester. So the amino group is deprotonated. It's got to be deprotonated to function as a general base catalyst, proposed to deprotonate the hydroxyl group of a threonine, activated for nucleophilic attack somehow. Do you form an oxyanion hole?

Whenever you see brackets, that means we don't see it and its a proposed intermediate. From chemical studies, we know tetrahedral intermediates exist. In proteases, no one has ever seen a tetrahedral intermediate. So all these things you see in all these mechanisms are a figment of people's imaginations based on really sort of a thorough understanding of the chemistry of what's going on. So when whenever you see brackets, that means there's no direct evidence or it's reasonable mechanism based on the chemistry.

And then what we need to do is we want to break down this tetrahedral intermediate. And we want to cleave the amide bond, which is the goal of this proteosome. And to do that, we can now use this amino group which we've initially used as a general base catalyst, as the general acid catalyst. So that's where the general acid catalysis comes in, and you see this over and over again in biology. You have one group and it can function as an acid and base catalyst.

And it gives you-- what does it give you? It gives you an acyl-enzyme. So you've seen that before. And now you just do the reverse of this reaction where this forms as a general base catalyst to activate water, forming a tetrahedral intermediate, and then loss of water to regenerate your starting material. So that's a working hypothesis and I don't really want to spend any more time on that. So that's the chemistry of the core particle.

And really sort of what we want to do now is focus on the real chemistry that's going to go on, that's going to allow us to mediate degradation of proteins of interest as a regulatory mechanism. So the second thing that I want to talk about, the second player I want to talk

about is ubiquitin, which you have all seen.

And this is the key tag that targets, although it's a major tag. But I think the tagging is much more complicated, as I tried to indicate in the first slide. We just went over it. You need something else to target your protein for degradation because you need to target the ubiquitination in the first place. So something has got to be special about protein A and protein D that controls lifetime inside the cell. And this is a lot of people working on that.

So tag the targets. Let's call it a protein of the interest. So this would be HMG-CoA reductase for degradation. So this is a key player and we're going to spend a little bit of time talking about that. But furthermore, one ubiquitin is not enough. You need to have polyubiquitins. So what we're going to see is-- and I think this is still the rule, it keeps changing-- but you need polyubiquitins where n is greater than or equal to 4 for this targeting to work, so one's not enough.

So now we're faced with the problem, how do we stick this ubiquitin onto the protein that's going to get degraded? So that's what we're focusing on. So let's look a little bit at the structure of ubiquitin. So here's the structure of ubiquitin, and ubiquitin is 76 amino acids. It's a compact-- you've already looked at the structure of this guy and it's compact a little protein that has a C-terminal glycine. Let me write this out.

So this is a C-terminal glycine, which is a key player. So in some of the pictures that you're going to see that I draw on the board, this is going to be a key player so I might write G76 or something like that, so glycine 76 is in all of these things. And it's going to be making the linkages we're going to be looking at. So you just need to remember that. And if you look at where it is, if you go back and you look at the structure, it's on a flexible loop at the end.

So the other thing that you need to know about ubiquitin when we look at the structure is that it has seven lysines. And the lysines, because this is so small, are located on the surface. For targeting ubiquitin, targeting proteins for degradation, we're really going to be focusing on, like, lysine 48. But all of the lysines can get modified. That's just the complexity of it. So we're going to focus on lysine 48.

So the key thing is what happens. What I'm going to show you is what the structure is. So we have a protein of interest that's targeted for degradation, and we're going to talk about how it's targeted. It has a lysine on its surface, a lysine on its surface, which is somehow then going to be attached to glycine 76 on the C-terminus of ubiquitin. So this is going to be part of ubiquitin.

So let me just write it-- I'm not going to write this out anymore, but this is a C-terminus. So what's unusual about this bond? I mean, do you normally see that bond in proteins? No. Normally, you see it with alpha amino groups, now you're seeing it with the epsilon amino group of lysine.

So this is an isopeptide and everything with ubiquitin in chemistry is through isopeptide linkages. So this is, again, this is a lysine here. And we will see that all the proteins targeted have lysines on the surface that can do covalent chemistry. And we'll see how with ubiquitin. So this is an isopeptide. Again, this is the epsilon amino group, not the alpha amino group. And this is the C-terminal flexible chain of ubiquitin.

And then, ultimately, what we need to get to when we're looking at the biosynthetic pathway is how do we attach all of these things? So the question is, what are the linkages between the ubiquitins? They're going to be isopeptide linkages, and they're isopeptide linkages between the lysine on the ubiquitin and the C-terminal glycine of another ubiquitin.

So what you have here again is an isopeptide. And I think once you get this down in your head as to what's going on, the chemistry is going to be really straightforward. So again, what you have in the case of ubiquitin, you have lysine. Remember, I told you we were going to focus on lysine 48, so lysine 48. Again, it is surface exposed and if forms a covalent linkage with this is glycine 76 at the C-terminus.

So again, we have this isopeptide linkage. So that you're going to see over and over and over again. Does everybody get that? I don't think it's so hard to see, but it's just different from what you've seen before.

So now I want to do-- I told you at the very beginning, how do we do all of that? Well, the time is over. I'm sorry. The time is over already. It just goes by so fast. I can't stand this. I didn't even get to the exciting part.

So next time, next time we'll have to come back and we will talk about E1, E2, E3. The chemistry is straightforward and it's analogous to chemistry you've already seen before. And then we're going to briefly look at how this relates to HMG-CoA reductase, degradation, which I'll show you what the factors are, but it's still pretty much a black box in my opinion.