ELIZABETH NOLAN: --by talking about ClpX. And then we're going to move into module 4-- which is the last module before spring break-- on synthases and assembly line biosynthesis.

So basically last time, where we left off is, we went over experiments that were done to look at denaturation, translocation, and degradation by ClpXP. And closed with a question about what actually is going on in ClpX with this ATP binding and hydrolysis to allow for these condemned protein substrates to be unfolded and translocated into the degradation chamber.

And I left you just with the statement that although we think about ClpX as this hexamer that has six identical subunits, what studies have shown is that there's some inherent asymmetry within this AAA+ ATPase. And that's what we're going to talk about a little bit.

So this is just a slide from a few lectures ago that's showing the top view and side view of ClpX and how we've thought about this. And I think these studies just really highlight how complicated these machines are and that there's still a lot more we need to figure out here.

So as I said last time, this asymmetry comes from whether or not each ClpX subunit is bound to nucleotide. And so basically, from looking at many different crystal structures, what can be done is that the ClpX subunits can be divided into two different types based on conformation here.

And so in thinking about this, we want to first think about the ClpX domain organization. And if we just think about this, what ClpX has is an end domain followed by a domain that's called the large domain and then followed by a small domain. So 633 amino acids, just to give you a sense of size, and about 69 kilodaltons per subunit.

And so what we're going to focus on are the large and the small subunits and what's observed from many different crystal structures. And so these two different types of subunit have been described as loadable and unloadable, and that depends on whether or not nucleotide is bound.

So if we consider of these two types, just thinking about the large and small domains, we have this loadable arrangement which binds nucleotide. And in cartoon, something like this. So we have the large subunit. We have the small subunit. And we have this region that's called a hinge. So this is one ClpX. So ATP binds.

And so the other type is described as unloadable. And this type of subunit does not bind nucleotide when in this unloadable conformation. And so we can draw this. Here, again, we have the large subunit. And there's a change in conformation. And here's the small subunit here.

So what's been found from looking at many crystal structures is that within the ClpX hexamer, there's an arrangement of these loadable and unloadable subunits. So in many crystals, what's found is that there's four loadable-- I'm just going to do with "L"-- plus two unloadable subunits arranged with about two-fold symmetry, so LULLUL. So there's some asymmetry in the subunits.

And so also from these crystal structures, there's some more observations that we don't see with just these cartoons of the 6-mer. So we can learn about how subunits interact, of course, and this is what's shown. So if we look at these structures and consider how these subunits interact, what we find is that the small subunit of one ClpX-- or sorry, the small domain of one ClpX subunit interacts with the large domain of the adjacent ClpX subunit.

And so we can draw this. Basically, if we consider a large domain-- and let's say this is subunit 2-- then what we find is that there's the small domain and then the large domain of subunit next door. So let's call this subunit 1. So here's our hinge of subunit 1, and ATP binding happens in here. So we can think about this arrangement.

And then what's been defined is something called a rigid body. And so this rigid body is comprised of the large domain of one subunit and the small domain of the next. Rigid body. So large domain of one ClpX and small domain of another subunit that's adjacent.

So in thinking about this, we can consider the ClpX hexamer in another way. So how I've initially presented it to you when we introduced these oligomers is just as a 6-mer, right? 6 subunits. But another way to think about ClpX is that it's actually six rigid bodies that are connected by hinges, where each rigid body has a component from two subunits, a large domain from one and a small from another.

And so the hinges are within a single subunit based on this cartoon where ATP binds. And so the thinking is that ATP binding and hydrolysis results in changes in the hinge geometry and that this change in confirmation in the hinge with ATP binding and hydrolysis allows for conformational change in another subunit here.

So six rigid bodies connected by six hinges, effectively, as opposed to just six standalone subunits. Each subunit's communicating with one another here. So this is pretty complicated, right? It's another level of sophistication within this hexamer here.

So what about these loadable and unloadable conformations? I've told you that in these crystal structures, what's seen often are these four loadable and two unloadable subunits with a particular arrangement. So we can ask the question, do these individual subunits maintain the same conformation during these attempts to denature and translocate polypeptides? So is one subunit just committed to being loadable and another subunit committed to being unloadable? Or did they switch dynamically?

And so recently, there were a number of studies looking at that. And effectively, as of a few years ago, many studies suggest-- or support switching by a given subunit. And they also indicate that every ClpX subunit must bind to ATP at some point during these cycles for unfolding and translocation. But they're not all doing it at the same time.

So the way to think about this is that there's some dynamic interconversion between these loadable and unloadable subunits within the hexamer, and somehow-- yep?

AUDIENCE: I just want to ask you--

ELIZABETH NOLAN: Going to make trouble?

JOANNE STUBBE: --a question. Yeah. So when you have all these structures, are they all with an ATP analogue?

ELIZABETH NOLAN: I don't know the answer to that.

JOANNE STUBBE: OK. Because ATP analogues have wide-- you guys have already seen ADPNP or ADPCH2P They really have very different properties when you study these, the ATPs. So if these-- and probably they don't have ATP because they probably--

ELIZABETH NOLAN: Right. They want to get a stable--

JOANNE STUBBE: So anyhow, that's something to keep in the back of your mind.

ELIZABETH NOLAN: So just is this an artifact is what JoAnne's suggesting from use of a non-hydrolysable ATP analogue.

JOANNE STUBBE: And there's many examples of that in the literature. Everybody uses it. It's just something you need to keep in the back of your mind. That's the best we can do.

ELIZABETH NOLAN: So next week, someone should ask during recitation there, for that.

So what about the mechanical work? How this is often depicted, in terms of grabbing and pulling on a polypeptide substrate, is via these rigid bodies. And we're not going to go into details about this, but just to describe the typical cartoon picture effectively, imagine we have some polypeptide that needs to enter the degradation chamber.

So those pore loops we heard about that are involved in substrate binding are in the large domain of ClpX. So here we have one large domain, and then we can have the small domain of the adjacent subunit here. And just imagine here we have another large domain with its pore loop. And then we'd have the adjacent subunit here.

So effectively, it's thought that these pore loops in the large domains grip the substrate and help drag the substrate to allow for translocation into the degradation chamber. So this would be going to the chamber, that direction here for that. So somehow the ATP binding and hydrolysis is allowing this to occur-- so to ClpP here for that.

So next week in recitation, you're going to have a real treat because an expert, Reuben, will be discussing some single molecule methods that have been applied to studying this degradation chamber. So bring your questions to him because he really knows what is state of the field right now for this.

So we've talked a lot about how the substrate needs to get in. We have the SSRI tag. We have all of this ATP consumption unfolding, translocation by ClpX. And then we talked about the serine protease mechanism in terms of how peptides are degraded in the chamber.

So then the final question just going to touch upon is, how does the polypeptide that's been degraded get out of the chamber? So ClpXP will give products that are 7 to 8 amino acids in length, so short polypeptides. So how are they released?

And we can think about two possibilities for how these polypeptides are released. One is that they're released through the axial pores. So somehow those pores that allow polypeptide substrate to go in also allow product fragments to go out. And then the second option is that there's release through transient side pores between the ClpP 7-mers.

So effectively, if we imagine coming back to our ClpP, we have a 7-mer-- back-to-back 7-mers, do the fragments come out, say, of the hole? Or somehow do they come out from this region here? To the best of my knowledge, this is a bit unclear, and I don't think they're mutually exclusive.

So questions have come up. If they're to come out of an axial pore, does that mean ClpX has to be dissociated? In terms of this equator region, there are structures showing that this degradation chamber can breathe. And there's a picture of that in the posted notes from Friday where you can see opening here. And there has been some experiments done where people have put cysteines in this region by site-directed mutagenesis.

So you can imagine, for instance, if you have a cysteine here and a cysteine here, and you oxidize to form a disulfide such that those two 7-mers are locked together. You can ask, if we load the chamber with small polypeptides and we have these effectively cross-linked by disulfides, can the polypeptides get out? And then if we reduce this to have them no longer attached to one another, do those polypeptides stay put or not there?

Those experiments gave some evidence for release of peptides through this region here, but there's also evidence for release of peptides through the pore. And in terms of cartoon depictions. In the lecture notes, if you take a close look, you'll see that both come out there. So I'd say if you're curious about that, you can read some of the literature and come to some own conclusions.

One last point on the Clp system before we move on to module 4, you should just be aware that there's other Clp family members. So not only ClpX and P. And so in the Clp system-- actually, I'm going to make one other point after this too, about degradation chambers.

So there are players ClpA, ClpB, in addition to ClpX. So these are all three different AAA+

ATPases. And you've actually encountered ClpB last week. So this is HSP 100, which came up in question 2 on the exam there by another name. And then in addition to ClpP, there's also ClpS and some other players here for that. They each have their own personality within protein quality control here for that.

And then we've only looked at this degradation chamber from bacteria. You might want to ask the question, what happens in other organisms? And the answer is that the complexity varies and systems become tremendously more complex as you move from bacteria into eukaryotes there.

And so if we consider the different degradation chambers, what do we see? So we find these proteasomes in all forms of life. And as I just said, the level of complexity varies depending on the organism. And so what we've seen with ClpP is the most simple system where we have two rings that have only one type of subunit. So just say E. coli. One type of subunit.

What happens if we go to archae? We find that we have four rings, each of which is 7-mer. And these four rings include two different types of subunits. So I'll call these alpha and beta. So what we find is that there's a 7-mer of 7 alpha subunits, then 7-mers that have 7 beta subunits, and here a 7-mer with alpha. So we see two types of subunit and four rings.

So then what about yeast? Tremendously complex. So we have this architecture again of four rings, an organized alpha, beta, beta, alpha. But what we find in this case is that in each of these-- I'm not going to draw it like that, but each of these have seven different subunits. There's a depiction of this in the notes.

So just imagine-- how does this get assembled? I have no clue. But somehow each of these heptamers has to be assembled with seven different subunits. And then they're put together in this series of four rings.

And then as you'll see after spring break in JoAnne's section, the eukaryotic proteasome has this 19S regulatory particle that's involved in recognizing condemned proteins that have polyubiquitin chains. And compared to the ClpX ATPase, it's much, much more complex.

So there's many different proteins that constitute this necessary part of the machine. But there is a hexamer, ATPase hexamer, within there to facilitate translocation of the polypeptide into the degradation chamber. So some of this will come back again in the latter half of the class.

So with that, we're going to close on degradation and move into module 4, which is focused on

macromolecular machines that are involved in the biosynthesis of natural products, specifically polyketides and nonribosomal peptides. And so we're completely taking a loop back to thinking about a biological polymerization, like what we were thinking with the ribosome from the process of breaking down a polypeptide.

And so where are we going? We can think about assembly lines, although this is a helpful way on the board to think about these systems. But it's not really what they look like. And you'll learn about that in recitation this week. Yeah?

AUDIENCE: Could you explain the interaction between ATP and the hinge area?

ELIZABETH NOLAN: OK. So the ATP binding site is just rewinding here in that hinge region. And there's going to be conformational change in the hinge with ATP binding and hydrolysis there. And that's sufficient in terms of the level of detail for this. But the main thing to keep in mind, each subunit binds ATP. But on the basis of the information gathered with the caveats JoAnne brought up, different subunits bind ATP at different times in the cycle.

AUDIENCE: OK. Thank you.

ELIZABETH NOLAN: And changes in this subunit, conformational changes that result from that, can be translated to the next door subunit here.

AUDIENCE: OK.

ELIZABETH NOLAN: OK. So where are we going? By a week from now, you should have a good handle on how to think about the biosynthesis of structures like erythromycin, of penicillin. These are products of assembly lines. And so where we'll go is with a brief overview of fatty acid biosynthesis and then look into polyketide synthase and nonribosomal peptide synthetase assembly lines here. And then some case studies.

So on the topic of ATP, where we just went back to with ClpX, just taking a look here, what do you know about these systems in ATP by the names? This is just a little language use and definition. So there's a subtle difference here. What's the difference?

AUDIENCE: Synthase versus synthetase?

ELIZABETH NOLAN: Yeah. And what does that tell you right off the bat? About ATP. So it's a subtlety, right? Synthase is a general term. Synthetase indicates ATP is involved. So as we'll see, these

nonribosomal peptide synthetases employ ATP. And we're going to see chemistry very similar to what you saw with the aminoacyl-tRNA synthetases in terms of activating amino acid monomers. But in this case, the machine is forming a nonribosomal peptide rather than a ribosomal peptide here.

If you are not familiar with fatty acid biosynthesis, I highly encourage you to go do some review, either from your 5.07 notes last term if you were in the class or from a biochemistry book. And there'll be some additional slides of overview information posted online. So we'll just touch upon it today but not go into tremendous detail here.

So what are our questions for this module? I think for most everyone in the room, this module will contain the most new information from the standpoint of a new system compared to what we've talked about so far. So what are polyketides and how are these molecules biosynthesized by polyketide synthases? What are nonribosomal peptides and how are they made by these machines called nonribosomal peptide synthetases?

And what we're going to look at is the assembly line organization, so effectively the organization of domains that provide these linear polymers. So what is the assembly line organization and logic for PKS? And likewise for NRPS.

And then we can ask, how can a given assembly line for a given PKS or NRPS natural product be basically predicted from the structure of the natural product? So you should be able to work back and forth in terms of looking at a structure and coming up with a biosynthetic prediction and also seeing biosynthetic machinery and getting a sense as to what that small molecule metabolite's backbone might look like. How are these studied experimentally? And we'll look at the biosynthesis of a molecule called enterobactin as a case study.

And so one thing I'll just point out right now is that these synthases and synthetases do not look like an assembly line. And we'll draw domains in a linear order which really facilitates thinking about the chemistry, but the structures are not just a line of domains or proteins next door to one another. And this week in recitation, you'll get to see some cryo-EM studies on fatty acid synthase and related machines there which will give you a sense of their dynamics.

So just a review. If we think about template-dependent polymerizations in biology, we're all familiar with DNA replication, transcription, and translation. And what you'll see in this unit is that these template-driven polymerizations occur in the biosynthesis of natural products here.

And effectively, these assembly lines, in a way, provide this template. So they're small molecules being biosynthesized by microbes using some pretty amazing machinery. So when we think about template-driven polymerizations, we think about an initiation process, elongation process, and termination. We saw that with a translation cycle. And we'll see the same type of systems here.

So what does some of these structures look like? Here are just some examples on the top of polyketides, two examples. They look very different at first glance, and they are. So we have tetracycline. We have four fused 6-membered rings. It's an aromatic polyketide, an antibiotic. We have this erythromycin here, which is a macrolide. We encountered macrolines in the translation section because they bind the ribosome, another type of antibiotic.

If we look at some nonribosomal peptides, all of that can be used clinically. We see the penicillins. So we have a 4 or 5 fused ring system here, a beta-lactam. This comes from three amino acid building blocks initially. We have vancomycin. This is an antibiotic of last resort.

And this structure looks really quite complicated, but what we'll see is that it's based on seven proteogenic amino acids. So it's the 7-mer peptide backbone that gives rise to this structure here for that. And then we see there's some sugars, so these can be put on by other enzymes here.

So on top we see a lot of ketones and OH groups. Those are good hints that maybe polyketide logic is being used. Here we see a number of peptide bonds, amide bonds, a good indicator of NRPS at play. And here, just to point out, these systems get very complex.

And there's natural products out there that are biosynthesized from a combination of polyketide synthase logic and nonribosomal peptide synthetase logic here. These include molecules like yersiniabactin. This is an iron chelator produced by Yersinia pestis, and some pathogenic E. coli, this immunosuppressant rapamycin as examples.

So as we move forward, I put a lot of structures of small molecule metabolites in the slides. You can go back and use them as a way to study and try to make predictions about what is the machinery at play, for instance, to give all of these heterocycles? How are those made? We'll see the assembly line does that.

So what organisms produce these molecules? Largely, bacteria and fungi. And there are some correlations out there, I'll just point out, related to genome size and the number of

metabolites being made. So bioinformatics guides a lot of current studies of the biosynthesis of these types of molecules.

So you can imagine that you sequence a genome. You have some information about gene clusters. So these are groups of genes where the proteins work together to biosynthesize the molecule. And often, the genes that encode proteins in these metabolites are clustered. And so bioinformatics approaches can help find these.

What's found is that for bacteria, some phyla are more prolific producers of these molecules than others. And what's been shown in a general way is that organisms with small genomes-- so something like E. coli-- produce fewer of these metabolites. That's not to say none. So enterobactin, which we'll look at for a case study, is made by E. coli. But they don't make as many.

And effectively, organisms with larger genomes produce more. And so here is just a correlation between the number of genes. And the genome size of the organism where they see around 3 Mb, there's a switch here. Often, these molecules-- yeah?

AUDIENCE: Is there any hypotheses about an evolutionary driving factor for the development of this machinery and why it correlates to genome size?

ELIZABETH NOLAN: If there is, I don't know. I don't think about evolution very well, quite frankly. What is thought is that many of these molecules are thought to be involved in defense and that an organism with a smaller genome size uses other strategies.

And so for instance, E. coli, which I cited as a small genome, will use a number of ribosomal peptides as defense molecules that get post-translationally modified after the fact. But why that organism chooses to do that versus say something like Streptomyces that produces many, many different natural products, I'm not sure about that.

So let's look at an example of a gene cluster, just so you get a sense of how much machinery is required to do the full biosynthesis of a molecule. So this is for a nonribosomal peptide shown here. It has some structural similarities to the vancomycin we saw on a prior slide, and it is a member of the vancomycin family.

So this gene cluster for the biosynthesis of this metabolite contains 30 different genes and is depicted here. So each one of these arrows indicates an open reading frame. So each one begins with a start, ends with a stop codon. And it's assumed to be the coding sequence of the

gene.

And so what is encoded in these 30 genes? Well, first there are the genes for what we call the assembly line. And if it isn't clear what assembly line means, as we move forward through this week, it will be. So there's genes required to make the 7-mer polypeptide backbone. There's genes required for modification of the backbone.

So how do these sugars get attached, for instance? Those are going to be some tailoring enzymes. And then if you take a close look, there's a number of non-proteinogenic amino acids in this molecule, and that means they have to come from somewhere.

And so this gene cluster also includes genes that are required for the biosynthesis of those monomers. So there's a lot of effort going in to making this molecule by some organism. And so presumably, under some set of conditions, it's important.

So moving towards the chemistry, with that background in hand, what are some points to make? So what we'll learn and see is that the assembly lines that produce the polyketides and nonribosomal peptides are macromolecular machines. So there's dedicated macromolecular machines for the biosynthesis of these secondary metabolites.

And so what are secondary metabolites versus a primary metabolite? So what's a primary metabolite?

AUDIENCE: I'm not even totally sure how to define metabolites. Isn't metabolites what goes in? Or what comes out?

ELIZABETH NOLAN: Rebecca?

AUDIENCE: Or easily produced directly from the materials the cell's consuming?

ELIZABETH NOLAN: So presumably, the cell needs to get materials to biosynthesize the secondary metabolites too, right? Somewhere, these amino acid monomers or the monomers that are used for polyketide synthetase need to-- they'd have to come from somewhere, right? So are primary metabolites important for growth?

AUDIENCE: Yes.

ELIZABETH NOLAN: Yes. Development? Reproduction?

AUDIENCE: Yes.

ELIZABETH NOLAN: Yeah, right. Under normal conditions, right? We're in trouble if we don't have our primary metabolites there, whether they're ingested or biosynthesized. What about a secondary metabolite? Just taking that--

AUDIENCE: I'm guessing it's not necessary.

AUDIENCE: --something we can make from primary metabolites?

ELIZABETH NOLAN: No. Well, you can. You can. So a secondary--

AUDIENCE: --necessary?

ELIZABETH NOLAN: Yeah. A secondary metabolite is not required for normal growth, development, reproduction. So for some reason, under some circumstances of need, these secondary metabolites get produced. So for some of these antibiotic molecules, maybe the organism needs to defend itself. In the case of enterobactin or yersiniabactin, maybe that organism needs iron. And so it's producing a molecule that will help it obtain that there.

So what is going on? We've seen some pretty complex molecules. What we're going to see is that these assembly lines convert simple acid monomers, if it's a polyketide synthase or amino acid monomers for a nonribosomal peptide synthetase, into linear polymers. So we're going to look at template-driven polymerizations that initially give linear polymers.

And in the case of PKS, this is very similar to fatty acid biosynthesis. What we see is that the assembly lines allow for iterative additions of malonyl and methylmalonyl units. And they catalyze carbon-carbon bond formations.

In the case of nonribosomal peptide synthetases, what we'll see is that these allow for condensations of amino acids to form peptide bonds and effectively form nonribosomal polypeptides. So polypeptide synthesis without the ribosome.

So even though the PKS and NRPS are forming a different type of bond and that requires different chemistry, what we'll see is that they use very similar logic. And just getting the logic

sorted out initially makes life much easier down the road. So take some time to look over the depictions in the notes outside of class as we go forward.

So these assembly lines use acyl or aminoacyl thioesters as the activated monomer units. So then how do we get from this linear polypeptide to some more complex structure? The short message on that is that the, quote, "polymers" that are produced-- and they may be short, right? We just saw-- they are short, 7 amino acids for vancomycin.

They can undergo further elaboration to give these complex structures. So there can be tailoring enzymes that work on the products of the assembly line. Or there can be domains in the assembly line that give additional activities that allow for methylation or cyclization here.

So we can think about fatty acid synthase as a paradigm here. And so if we think about fatty acid biosynthesis making some molecule like this oil here, just as brief overview in the last few minutes of class. Fatty acids are synthesized by FAS.

And what happens is that there's elongation by one unit at a time. And each unit provides two carbons. So there's two carbon atoms per elongation. And so hopefully you're all familiar with two ways to form a carbon-carbon bond, at least related to biochemistry, one of which is Claisen condensations. So Claisen condensations allow for carbon-carbon bond formation and join the units.

To keep in mind, the monomers are always thioesters, not oxoesters. And for fatty acid biosynthesis, the two monomer units are shown here. So we have a starter and an extender, acetyl CoA or malonyl CoA here. And here we have coenzyme A.

So just as a brief review, if we think about these monomer units-- so here we have acetyl CoA. So what can we say about this guy here, in this thioester? So is this acidic or not? Compared to an oxoester. How many of you have heard about fatty acid biosynthesis?

AUDIENCE: [INAUDIBLE]

ELIZABETH NOLAN: So why are thioesters used and not oxoesters?

AUDIENCE: [INAUDIBLE] use the other end?

ELIZABETH OK. So we'll go into a little more detail on Friday to make sure the chemistry is straight here

NOLAN: because I'm not certain it is. So--

AUDIENCE: Is oxoester referring to not that [INAUDIBLE]--

ELIZABETH NOLAN: OK. So for Friday, think about a thioester versus an oxoester, and how do properties differ? And why might we want to be using thioesters? And also review the Claisen condensation because that's the chemistry that's going to be happening to form the carbon-carbon bonds in the fatty acid synthase and in the polyketide synthases.

And what we're going to see is that the monomers in each case, they're tethered as thioesters. So why is that? And I will turn around and point at somebody, and you can let us know. Are you excited? OK. So you're off the hook for Wednesday. I need to be out of town, and I'll see you on Friday.