ELIZABETH NOLAN: We're going to get started and what we'll do today is continue with fatty acid synthase. Because that's the paradigm for these macromolecular machines, like the PKS, and then we'll go over the logic of polyketide synthases. So we left off last time with this discussion about some molecules that will be involved and in particular thioesters, and I asked about the alpha H. So just going back to introductory organic chemistry, what are the properties of this atom here?

AUDIENCE: [INAUDIBLE] acidic.

ELIZABETH NOLAN: Yeah. OK, right. So this is acidic. So if you have-- OK? So what that means is if there is a base that can deprotonate that, we can get an enolate. OK, and this is the type of chemistry that's going to be happening with the thioesters that are used in fatty acid synthase and also polyketide synthase.

And just to rewind a little bit more, if we think about carbon-carbon bond forming reactions in nature, which is what's happening in fatty acid biosynthesis and in polyketide biosynthesis, effectively, nature uses three different types of reaction. OK, so one is the aldol, two are the Claisen, and three [INAUDIBLE] transfer. OK, and so we're going to see Claisen condensations in FAS and PKS biosynthesis. And then after spring break, when Joanne starts with cholesterol biosynthesis, that will involve [INAUDIBLE] transfers. And hopefully, you've seen aldol reactions sometime before within biochemistry here. OK?

So we need to think about just what the general Claisen condensation is that we're going to be seeing here and the consequences of this acidic proton. So also just keep in mind, rewinding a little more, nature uses thioesters not esters, and so the alpha H is more acidic. The carbonyl is more activated for nucloephilic attack. And there's some resonance arguments and orbital overlap arguments that can guide those conclusions, if you wish to do them here.

OK. So let's imagine that we have a thioester. We have a base. OK, that's going to be [INAUDIBLE], which is going to get us to here. So this is our nucleophile, and what you'll see

coming forward is an enolate.

So imagine we have that, and we add it with another thioester, and here's our electrophile. What do we get? We get formation of a beta-keto thioester, which is the Claisen condensation product. OK, you have two thioesters. OK?

So effectively, this acyl thioester is doubly activated, so it can be-- did I lose it? Oh no, problems. Sorry about that. It can be activated as an electrophile at the C1 position, so next door to the sulfur. And it can be activated as a nucleophile at the C2 position here.

So this is the general chemistry that's going to be happening by FAS and PKS in terms of forming carbon-carbon bonds between monomers here. OK? So in fatty acid synthase, we have two monomer units. OK? So we have acetyl-CoA and malonyl-CoA. Acetyl-CoA is the starter unit, sometimes called unit 0, and then malonyl-CoA is the extender.

And so recall that in fatty acid biosynthesis, each elongation event adds two carbons, and if we look at malonyl-CoA, we have three here. Right? So there's decarboxylation of malonyl-CoA to generate a C2 unit, and there's details of that in the lecture 15 notes. And SCoA is coenzyme A, here, and there's some information as to the biosynthesis of these starter and extender units in the notes. We're not going to go over that in lecture here.

So in terms of using these monomers to obtain fatty acids, first what we're going to go over are the domains in FAS. And so we can consider domains that are required for extension of the fatty acid chain and then domains that are required for tailoring of that effectively to reduce the carbonyl, as shown. And we're going to go through these, because what we're going to find is that with polyketide biosynthesis, the same types of domains are used. So this logic extends there.

OK. So first, we have domains required for elongation of the fatty acid chain by one two-carbon unit. OK. So these include domains that may be abbreviated as AAT or MAT, and they can be grouped as AT and stand for acetyl or malonyltransferase. OK. We have an Acyl Carrier Protein, ACP, and this carries the growing chain between the domains of fatty acid synthase. And so in recitation this week, you're going to see how these domains move around and talk about the length of this acyl carrier protein.

We also have the ketosynthase. So what the ketosynthase does is it accepts the growing chain from the acyl carrier protein, and it catalyzes the Claisen condensation with the next monomer.

And what we'll see is that this ketosynthase uses covalent catalysis, and via a cysteine thiolate residue. So these are the key domains required for elongation of the chain. OK?

And then what we also need are domains required for tailoring, and just to clarify, I'm defining domain here as a polypeptide with a single enzymatic activity. So domains can be connected to one another, or they can be standalone in different types of synthases, but domain means polypeptide with a single enzymatic activity. So what are the domains required for tailoring? And these work after addition of the C2 unit to the growing chain.

So first, there's a ketoreductase. And as indicated, what this enzyme does is it reduces the carbonyl of the previous unit to an OH and uses an NADPH H plus. We also have the dehydratase here, and this forms an alpha, beta-alkene from the product of the ketoreductase action.

And then we have an enoyl reductase that reduces this alpha, beta-alkene, and this also requires NADPH H plus here. And then some fatty synthases use a domain called a thioesterase for chain release, and that's noted as TE. And we'll see thioesterases in the PKS and in our PS sections here. So one comment regarding the acyl carrier protein, and then we'll just look at the fatty acid synthase cycle and see how these domains are acting.

So in order for the acyl carrier protein to carry this growing chain, it first needs to be post-translationally modified with what's called a PPant arm. And that arm provides the ability to have these monomers, or growing chains, linked via a thioester. And so just to go over this post-translational modification, so post-translational modification of acyl carrier protein with the PPant arm.

OK. If we consider apo acyl carrier protein, and apo means that the PPant arm is not attached. There's a serine residue. An enzyme called the PPTase comes along, and it allows for post-translational modification of this serine using CoASH, releasing 3', 5'-ADP to give ACP post-translationally modified with the PPant arm. OK?

And we'll look at the actual chemical structures in a minute. What I want to point out is that throughout this unit, this squiggle, some form of squiggle here, is the abbreviation for the PPant arm. OK? And this is very flexible and about 20 angstroms in length.

So what does this actually look like? So here we have CoASH. So PPant is an abbreviation for phosphopantetheine, here, this moiety, and here's the 3', 5'-ADP. And so effectively, what's

shown on the board is repeated here. Except for here, we're seeing the full structure of the phosphopantetheinylated acyl carrier protein.

So this squiggle abbreviation indicates this post-translational modification onto a serine residue of the ACP. Just as an example of structure, so here is a structure of acyl carrier protein from E. coli. It's about 10 kilodaltons, so not very big, and we see the PPant arm here attached. OK?

So if we think about fatty acids biosynthesis, we can think about this in three steps, better iterated. OK. So first we have loading, so the acyl carrier proteins need to be loaded with monomers. Sometimes, this step the reactions are described as priming reactions. We have initiation and elongation all grouped together here and, three, at some point, a termination. OK? So we've thought about these before from the standpoint of biological polymerizations.

So what about the FAS cycle? Here's one depiction, and I've provided multiple depictions in the lecture 15 notes. Because some people find different cycles easier than others, but let's just take a look.

So this charts out the various domains-- the starter and the extender and then the chemistry that occurs on these steps. And so what needs to happen is that there needs to be some loading and initiation where the acetyl-CoA is loaded onto an acyl carrier protein. So that's shown here via transferase here, and then, from the acyl carrier protein, this monomer is loaded onto the ketosynthase.

If we look here, we have one of our extender units, the malonyl-CoA, and the CO2 unit that gets removed during decarboxylation, as shown in this light blue. OK? We need to have this extender unit also transferred to an acyl carrier protein via the action of an AT. So we see lots of the CoA. Here we have the acyl carrier protein with the PPant arm. It's not a squiggle here. It is the next one with this malonyl unit loaded.

There's a decarboxylation, and what do we see happening here? We have a chain elongation event, so Claisen condensation catalyzed by the ketosynthase between the starter and the first extender to give us this beta-keto thioester. So once this carbon-carbon bond is formed to give us the beta-keto thioester, there's processing of the beta carbon via those tailoring domains-- the dehydratase and the enoyl reductase. And so we see reduction of the beta ketone here, we see formation of the alkene, and then we see reduction to get us to this point.

And so this cycle can repeat itself until, at some point, there's a termination event. And in this case here, we see a thioesterase catalyzing hydrolytic release of the fatty acid chain. This is the depiction you'll see in recitation today, or saw before. And I guess what I like about this depiction is that you see color coding separating the elongation and the domains involved in elongation with then the processing of the beta ketone here and then termination. OK. So we get some fatty acid from this.

And so where we're going to go with this overview is looking at the polyketides and to ask what similar and different in terms of polyketide biosynthesis? And so where we can begin with thinking about that is asking what are the starters and extenders? And so these are the starters and extenders we saw for fatty acids, and here are the starters and extenders for polyketides, so very similar. Right? We just see that there's some additional options, so we also have this propionyl-CoA here. In addition to malonyl-CoA as an extender, we see that methylmalonyl-CoA can be employed.

So what are the core domains of the PKS? They're similar to those of FAS, and we'll just focus on the PKS side of this table. So this is a helpful table when reviewing both types of assembly lines.

So the core means that every module, which I'll define in a moment, contains these domains. So we see that there's a ketosynthase, an acyltransferase, and a thiolation domain. So this thiolation domain is the same as the acyl carrier protein. So there's different terminology used, and within the notes, I have some pages that are dedicated to these terminologies. OK? So for PKS, here, we have the ketosynthase, we have acetyltransferase, and then we have this T domain which equals acyl carrier protein here. OK?

So then what about these tailoring domains that were required to produce the fatty acid? What we see in polyketide biosynthesis is that those domains are optional. So one or more of these domains may be in a given module.

So that's an overview, and then we'll look at an example of some domains and modules. So we're going to focus on type 1 polyketide synthases. And in these, what we're going to see is that catalytic and carrier protein domains are fused, and they're organized into what we'll term modules. So a module is defined as a group of domains that's responsible for activating, forming the carbon-carbon bonds and tailoring a monomer.

So there is an individual module for every monomer within the growing chain. And the order of

the modules in the polyketide synthase determines the functional group status, and that functional group status is determined by whether or not these optional domains are there. OK? How do we look for modules? The easiest way is to look for one of these thiolation or ACP domains. So each module has one of these. So you can count your number of T domains, and then you know, OK, there's 7T domains, so there's 7 monomers, for instance.

So each Claisen condensation is a chain elongation and chain translocation event. Keep in mind, the starting monomer-- so whether that's acetyl-CoA or propionyl-CoA-- does not contain a CO2 group. So there's no decarboxylation of the starting monomer, but decarboxylation of malonyl-CoA occurs, like in fatty acid synthase, and if that's the case, it provides a C2 unit. And if methylmalonyl-CoA is the extender, this decarboxylation provides a C3 unit because of that methyl group.

So key difference, as we just saw, in fatty acid biosynthesis, we have complete reduction of that beta-keto group in every elongation cycle because of these three tailoring domains-- the KR, DH, and ER. In PKS, what can happen is that reduction of this beta-keto group may not happen at all, or it may be incomplete in each elongation step. So what that means is that polyketides retain functional groups during chain elongation. And if you look back at some of the structures that were in the notes from last time, you'll see that, in terms of ketones, hydroxyls, double bonds, et cetera.

And also, the other point to note is that there can be additional chemistry, and that these assembly lines where polyketide synthases, non-ribosomal peptide synthatases can contain what are called optional domains. So these are additional domains that are not required for formation of the carbon-carbon bond or amide bond in non-ribosomal peptide synthases. But they can do other chemistry there, so imagine a methyltransferase, for instance, or some cyclization domain.

So how do we show these domains and modules? So typically, a given synthase is depicted from left to right in order of domain and bond-forming reactions here. So let's just take a look. So if we consider PKS domains and modules, we're just going to look at a pretend assembly line. OK? So this I'm defining here as an optional domain.

So in this depiction, going from left to right, each one of these circles is a domain, so a polypeptide with a single enzymatic activity. Note that they're all basically touching one another which indicates in these types of notations that the polypeptide continues. It's not two different

proteins, but we have one polypeptide here.

I said that there's modules, and we can identify modules by counting T domains. So here, we have three T domains. So effectively there's three modules. So we have a module here, we have a module here, and we have a module here.

What do we see? Two of these modules have a ketosynthase, so that's the domain that catalyzes the Claisen condensation. We have no ketosynthase here, in this first module.

Why is that? We're all the way to the left. This is effectively our starter or loading module. So the propionyl-CoA or acetyl-CoA will be here, as we'll see, and there's nothing upstream to catalyze a condensate event with. So there's no KS domain in the starting module here or loading module. OK.

So this is often called loading or starter. So if we think about these optional domains for a minute and think about how they work. If we go back to fatty acid synthase, and let's just imagine we have this species attached. We have the action of the KR, the dehydratase, and the ER to give us the fully-reduced species. Where here, we have a CH2 to group rather than the beta-ketone.

So what happens in PKS in terms of the different optional domains? So we could have this and have full reduction. We can imagine maybe there's no enoyl reductase. So the module has the ketoreductase and the dehydratase but no enoyl reductase, and so as a result, this polyketide ends up with a double bond here. OK?

What if we have nobody dehydratase, like this? OK. We just work backwards from the FAS cycle. We'd be left with this OH group at the beta position. Right? And if we have none of them, so no ketoreductase, dehydratase, or enoyl reductase, the beta-ketone will be retained, here. So what this also means is that you can just look at some polyketide and assess what the situation is from the standpoint of these optional domains.

So let's just take an example. If we have three cycles of elongation, and let's imagine we had an acetyl-CoA starter plus three malonyl-CoA. So what do we end up with? Let's imagine our chain looks like this.

What do we see? So two carbons are added during each elongation cycle to the chain here, and we can see those here, here, here, and here. OK? So a total of four C2 units, one from the starter and then three from these three extenders. And then we can look at what the

functional group status is and say, OK, well here, we have no ketoreductase. And here, there was ketoreductase action, but there's no dehydratase.

And here, what do we see? We see that there was a reduction of the beta-ketone and then the action of the dehydratase, but we're left at the alkene, so no enoyl reductase. Right? So just looking, you can begin to decipher in a given module what optional domains are there.

So what we'll do is take a look at an actual PKS assembly line and then look at the chemistry happening on it here. These are just for your review. This is a polyketide synthase responsible for making this molecule here. So D-E-B or DEB is a 14-membered macrolactone. It's a precursor to the antibiotic erythromycin here, and this is the cartoon depiction of the polyketide synthase required for the biosynthesis of this molecule.

So what do we see looking at this polyketide synthase? So it's more complicated than this one here, but the same principles apply. And what we'll see is that it's comprised of three proteins. There's seven modules, so one loading or starter module and six elongation modules, and there's a total of 28 domains. OK? And I said before, the placement and the identity of these domains dictates the identity of the growing chain.

So let's take a look. So first, how do we know there's three proteins? We know that in this type of cartoon because we end up seeing some breaks between different domains. So here, for instance, the AT, the T, the KS, et cetera, they're all attached to one another in the cartoon. That means it's all one polypeptide chain, but this one polypeptide chain has many different enzymatic activities in it, because it has different domains.

When we see a break-- so for instance here this T domain and this KS domain are not touching one another. That means we have two separate proteins. So this T domain is at the terminus of DEBS 1, and DEBS 2 begins with this ketosynthase. OK?

Likewise, we have a break here, between the T domain and this ketosynthase. So three proteins make up this assembly line, and so when thinking about this, these proteins are going to have to interact with each other in one way or another. And so there's a lot of dynamics in protein-protein interactions happening here.

How do we know there's seven modules? And remember each module is responsible for one monomer unit. We count the T domains, so we have one, two, three, four, five, six, seven T domains. So like the acyl carrier proteins of fatty acid synthase, these T domains will be post-

translationally modified with a PPant arm. And that PPant arm will be loaded with the acetyl-CoA or methylmalonyl-CoA or malonyl-CoA monomers.

We have a loading module. So the loading module has no ketosynthase, because there's nothing upstream over here for catalyzing a carbon-carbon bond formation event. And then we see modules one through six, so sometimes the loading module is module zero. We see that each one has a ketosynthase, so there'll be carbon-carbon bond formation going along this assembly line.

And we see that the optional domains vary. So for instance, module one has a ketoreductase as does module two. Look at module four. We see all three domains required for complete processing of that beta-keto group here. Here, only a ketoreductase, and here only a ketoreductase. OK?

So just looking at this, you can say, OK well, we'll have an OH group here, here. Here we have complete processing. Just ignore this. It's in lower case, because it's a non-functional reductase domain. It's not operating as annotated here.

So what happens? So again, there's post-translational modification of this T domain, so it has a serine. The serine gets modified with the PPant arm, as shown here, and we use that squiggle depiction, as I showed for the acyl carrier protein of FAS. So post-translational modification of these T domains has to happen before any of the monomers are loaded onto this assembly line. And these PPant arms allow us to use bioesters as the linkages and through the chemistry I showed earlier.

So here, what we're seeing in this cartoon, going from here, this indicates that the T domains are not post-translationally modified. And here, we see the assembly line after action of some [? phosphopentyltransferase ?] loading these arms. OK? So each T domain gets post-translationally modified.

What happens next? We have loading of monomers. And we'll look at module zero and one on the board and then look at how the whole assembly line goes.

AUDIENCE: Do you ever get selected post-translational modification of the T domains and if so, does that facilitate different modules being like on or off, so to speak?

ELIZABETH I don't know. I don't know in terms of the kinetics, and say, does one T domain get loaded by a

NOLAN: PPTase before the other? These enzymes are very complex, and there's a lot we don't know. But that would be interesting, if it's the case. I wouldn't rule it out, but I just don't know.

One thing to point out too, these assembly lines are huge. So this is something we'll talk about more the next time, as we begin to discuss how do you experimentally study them? But some are the size of the ribosome for the biosynthesis of one natural product. And what that means, from the standpoint of in vitro characterization, is that often you just can't express a whole assembly line, let alone say one protein that has a few modules.

So often, what people will do is individually express domains or dye domains and study the reactions they catalyze in their chemistry there. And so it would be very difficult even to test that in terms of in vitro. Is there an ordering to how the T domains are loaded? And then there's question too, do you even know what the dedicated PPTase is? So there's some tricks that are done on the bench top to get around not knowing that, which we'll talk about later.

So back to this assembly line to make DEB. So we're just going to go over the loading module and module 1 and look at a Claisen condensation catalyzed by the KS. And this chemistry pertains to the various other modules and other PKS. So we have our AT domain and our thiolation domain of module 0, and then we have the ketosynthase, the AT domain, the ketoreductase, and the T domain of module 1. OK.

I'm drawing these a little up and down just to make it easier to show the chemistry. So sometimes you see them straight, sometimes moved around here, but it's all the same. So we have these PPant arms on the two T domains.

So what happens now, after these have been post-translationally modified? We need the action of the AT domains to load the monomers onto the PPant arms here, so action of the AT domain. So what do we end up with? In this case, the starter is a propionyl-CoA, so we can see that here. And we have a methylmalonyl-CoA as the extender, that gets loaded, and I'm going to draw the cysteine thiolate of the ketosynthase here.

So what happens next? We need to have decarboxylation of the methylmalonyl-CoA monomer to give us a C3 unit. And it's C3 because of this methyl group, but the growing chain will grow by two carbons. And then we need to have transfer of this starter to the ketosynthase.

So the ketosynthase is involved in covalent catalysis here. So what happens, we can imagine here, we have attack, and then here, we're going to have the decarboxylation. We have chain

transfer to the ketosynthase, and here, decarboxylation leaves us this species. OK? OK.

So now, what happens? Now, the assembly's set up for the Claisen condensation to occur which is catalyzed by the ketosynthase. Right? So what will happen here? You can imagine that, and as a result, where do we end up? I'll just draw it down here.

And what else do we have? We have a ketoreductase. So this ketoreductase will act on the monomer of the upstream unit, and that's how it always is.

So if there's optional domains in module 1, they act on the monomer from module 0. If there's optional domains in module 2, they'll act on the monomer for module 1. OK? So we see here now we have reduction of the ketone from module 1 to here via the ketoreductase. OK?

So if we take a look at what's on the PowerPoint here, what we're seeing is one depiction of this assembly line to make DEB indicating the growing chain. OK? So as we walk through each module, we see an additional monomer attached. So the chain elongates, and then you can track what's happening to the ketone group of the upstream monomer on the basis of the optional domains here.

If we look in this one, which I like this one because they color code. So they color code the different modules along with the monomer, and so it's pretty easy to trace what's happening. So for instance, here we have the loading module, and we have the starter unit in red. And here we see that it's been reduced by the ketoreductase of the upstream blue module. Here, we have the green module, here is its monomer, and we see its ketoreductase acted on the blue monomer from module 1, et cetera here.

So I encourage you all to just very systematically work through the assembly lines that are provided in these notes, and it's the same type of chemistry over and over again. And if you learn the patterns, it ends up being quite easy to work through, at least the simple assembly line. So as you can imagine, complexity increases, and we'll look at some examples of more complex ones as well. So where we'll start next time with this is just briefly looking at chain release by the thioesterase. And then we'll do an overview of non-ribosomal peptide biosynthesis logic and then look at some example assembly lines.

So we have the exams to give back. I'll just say a few things. So the average was around a 68, plus or minus 10, 11, 12 for the standard deviation. I'd say, if you were in the mid 70s and above, you did really well. If you're into the low 60s, that is OK, but we'd really like things to

improve for the next one.

In terms of the exam and just some feedback-- and I'll put feedback as well in the key which will be posted later today or early tomorrow. There wasn't one question that say the whole class bombed, so that's good. There were a few things for just general improvement, and I want to bring this up, so you can also think about it in terms of problem sets. One involves being quantitative. So there's certainly qualitative trends and data, but there's also quantitative information there, and that can be important to look at.

And one example I'll give of that involved question one. If you recall, there was an analysis of GDP hydrolysis and an analysis of peptide bond formation. And quantitative analysis of the peptide bond formation experiments will show that all of the lysyl-tRNAs were used up in the case of the codon that was AAA. Whereas, some of those tRNAs were not used up when the codon contained that 6-methyl-A in position one. Right?

And if you linked that back to the kinetic model along with the other data, what that indicates is that proofreading is going on. Right? Some of those tRNAs are being rejected from the ribosome there. So that was one place where quantitiation, a fair number of you missed that.

And another thing I just want to stress is to make sure you answered the question being asked. And where an example of that came up was in question one with the final question asking about relating the data back to the kinetic model. And so if a question asks that you really do need to go back to the model which was in the appendix and think about that.

So many of you gave some very interesting answers and presented hypotheses about perhaps the 6-methyl-A is involved in regulation and controlling like the timing of translation. And that's terrific and interesting to think about, but it wasn't the answer to the question. Right? Which was to go beyond the conclusions from the experiments with GTP hydrolysis and formation of that dipeptide, and ask how can we conceptualize this from the standpoint of the model we studied in class?

And then just the third point I'll make is related to question two and specifically to GroEL. But the more general thing is that if we learn about a system in class, unless there's compelling data presented in a question to suggest the model is something other than what we learned or its behavior is something other than what we learned, stick with what you know. So in the use of GroEL, the idea in that experiment was that, if you recall, this question was looking at these J proteins and asking, how do J proteins facilitate disaggregation? Right?

And so a GroEL trap was used that cannot hydrolyze ATP, which means it's not active at folding any polypeptide. But the idea there is that these J proteins end up allowing monomers to come out of the aggregate, and then GroEL can trap and unfold the monomer to prevent reactivation. And so a number of people came to the conclusion that GroEL was binding that aggregate somehow in its chamber. And what we learned about GroEL is that its chamber can't house a protein over 60 kilodaltons. Right? We saw that in terms of the in vitro assays that were done looking at what its native substrates are. Right?

So always go back to what you know, and then you need to ask yourselves, are the data suggesting some other behavior? And if that were the case, like what is your analysis of those data there? So please, even if you did really well, look at the key and see what the key has to say. And if you have questions, you can make an appointment with me or come to office hours or discuss with Shiva there. OK?