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

ELIZABETH NOLAN:
So where we're going to begin today is continuing with our discussions of the substrates for groEL, groES, and analysis of the data. And after that we'll talk about the DNAK DNAJ chaperone system here. So recall last time we left off with the question of the groEL groES substrate. So inside an E coli cell, what are the polypeptides that are folded by this macromolecular machine?

And so there was the pulse chase experiment, there was immuno precipitation, and then analysis. And so in this analysis, we talked about doing two dimensional gel electrophoresis, and then trypsin digest and mass spec of the various spots. So where we left off were with these data here and the question, how many polypeptide substrates interact with groEL in vivo, so inside an E coli cell?

And what we're looking at are the various gels for either total soluble cytoplasmic proteins on top at either 0 minutes-- so at the start of the pulse-- recall that these cells were treated with radio labeled methionine, and then there was a chase for a period of time when excess unlabeled methionine was added. So here we're looking at total soluble cytoplasm proteins 10 minutes into the chase. And then at the bottom, what we're looking at are the polypeptides that were immunoprecipitated by treatment of this cell lisate say with the anti groEL antibody. So the idea is this antibody will bind to groEL, and if polypeptides are bound those will be pulled down as well.

So it's kind of incredible this experiment worked. There was a bunch of questions after class in terms of the details of this immunoprecipitation just to think about is it a groEL monomer, or is it a groEL heptamer? How tightly are these polypeptides bound? How do they stay bound during the course of the workup? Where's groES? These are a number of questions to think about and to look at the experimental to see about answers.

So where we're going to focus is right now looking at these gels. And so what we need to ask is, what do we learn just from qualitative inspection of these data? So on these along the y-

axis we have molecular weight, and along the x-axis the PI. So if we first take a look at the total soluble cytoplasmic proteins at zero minutes and 10 minutes, what do we see? Do we see many spots or a few spots? Many spots, right?

And we see many spots both at 0 minutes and at 10 minutes. So the E coli genome encodes over 4,000 proteins-- roughly 4,300. And if one were to go and count all of these spots, how many do we see? It's on the order of 2,500. So they detected on the order of 2,500 different cytoplasmic proteins on these gels.

What do we see in terms of distribution by molecular weight? Is it a broad distribution, or narrow distribution? Broad, we're seeing spots of all different molecular weights, so from low to high on this gel. What about PI?

AUDIENCE: It's also broad.

ELIZABETH NOLAN: We also have a broad distribution in these gels, right? So we see polypeptides of low through high PI on this scale from 4 to 7. So now what we want to do is look at the gels obtained for the samples from the immunoprecipitation and ask what do we see, and is that the same or different from what we see for the total cytoplasmic proteins up here?

So if we look at the data here which are the polypeptides that were obtained from immunoprecipitation at 0 minutes, what do we see? So do we see a few spots, a lot of spots?

AUDIENCE: It's still a lot, and it's still distributed over a pretty wide range.

ELIZABETH NOLAN: OK, so let's start with the first point Kenny made, which is that we have a lot of spots, and I'd argue that's true. In this gel, we see many spots where each spot indicates a distinct polypeptide. Do we see the same or less than here for the total cytoplasmic protein?

AUDIENCE: It's less.

ELIZABETH NOLAN: We see less, right?

AUDIENCE: And they seem more concentrated.

ELIZABETH NOLAN: Yeah, just wait a second. Right, so we see less, and that's a good sign because an antibody was used to pull down some fraction of this pool. So about how many are here? They found about 250 to 300 polypeptides there, so about 10% of these cytoplasmic proteins were found

to be interacting here.

So on the basis of the experiment, we can conclude these are polypeptides that interact with groEL here. OK so now Kenny has a few additional observations in this gel. What are those? So how are these polypeptides distributed? And we'll just focus on C for the moment. So in terms of molecular weight, what do we see?

AUDIENCE: It's all scattered pretty wide range of molecular weights.

ELIZABETH NOLAN: And so we have a wide range, and where is that range and how does that range compare to here? So I agree, but look at the subtleties.

AUDIENCE: Most of them are above 8 kilodaltons?

ELIZABETH NOLAN: Yeah, so let's roughly say in the range of 20. So if we look at the bottom part of the gel versus the top part of the gel here, and we compare that to the bottom part of the gel here and the top part of the gel here, we see some differences that aren't just the total number of spots. Rebecca?

AUDIENCE: So it's like the ones that are smaller-- so the spots that respond to the smaller proteins, they seem to be more highly charged.

ELIZABETH NOLAN: More highly charged. Yeah, so let's first stick to the size. So we're seeing that in the bottom region of this gel where we have lower molecular weight species, we see fewer of these here than here. So why might that be if there's less polypeptides with molecular weight smaller than 20 kilodaltons? Steve?

AUDIENCE: If you just consider the total number of possible confirmations of protein can adopt or peptide to adopt as a exponential function of its size, larger proteins are more likely to have more non-productive folding pathways. So it's just less likely to have something that needs a chaperone at a smaller size.

ELIZABETH NOLAN: Right, so maybe these smaller polypeptides, they need less help. Their domain structure is more simple. For instance, they're easier to fold, and other machinery can take care of that here. And then if we look at PI, what do we see? So how is the distribution in terms of PI?

AUDIENCE: Large molecular weight proteins are pretty evenly distributed, but the smaller ones have more of a charge.

ELIZABETH NOLAN: Yeah. How do you use the word charged?

AUDIENCE: Sorry, I was looking at the scale. They actually have a PI closer to 7.

ELIZABETH NOLAN: Yeah, just like you heard in recitations 2 and 3, pay attention to the scale and what kind of charge-- if you're talking about charge, you have negatively charged and positively charged amino acids. So where in that regime are you? But if we look at these areas here, we see a wide distribution. And maybe when they're smaller we're seeing some more over here, but then ask yourself, is 22 an outlier there?

So what can be done in terms of these data? This is actually an analysis of the gels looking at total proteins and groEL bound proteins for the total percentage in terms of PI and in terms of molecular weight. And so you can compare. And so what we see is that overall, and look a bit closer, that PI distributions are quite similar. Molecular weight we see some differences.

We also don't see that many proteins that are greater than 90 kilodaltons being folded by this machine. And then again, why might that be? We learn that the chamber can accommodate polypeptides up to about 60 kilodaltons, so maybe they're just too big here.

So what are the identities of these proteins here? So this is where the trypsin digest and mass spec comes into play. So you can imagine extracting the spots, digesting them with the protease trypsin, and then doing mass spec analysis to find out the identities and comparing that data to databases of E coli proteins.

And so from that, of the 250 to 300 proteins that they identified in these immuno precipitation gels, they were able to identify 52 without a doubt. And what are some of those 52 proteins? So I've just highlighted a few examples. What do we see?

So here's our friend DFTU as one example. We see subunit of RNA polymerase, ferritin, and certain rhibosomal proteins. So just thinking about these proteins and their role in translation, in RNA polarization, ferritin is an iron storage protein. What do we think? What are our thoughts about these proteins?

They're pretty important, right? Imagine if EFTU you couldn't adopt its native confirmation. There might be some major problems. And recall when I introduced groEL, groES, we learned that they fall into the category of chaperonin, so they're essential for life. So that makes sense

in terms of seeing some of these proteins as being very important.

And what about structural motifs? It's then we see, OK, these are the 50 proteins we identified, what are their structural features, and what does that tell us about this chaperone? The conclusion is that overall, the proteins identified have quite complex structural features. So these can range from complex domain organization to beta sheets, including those that are buried and have large hydrophobic surfaces here.

And so we can speculate that maybe some of these hydrophobic surfaces interact with the groEL-applicable domain to have these polypeptides enter into the chamber. Here, was there a question?

AUDIENCE: Well, I was going to ask, I don't know for ferritin, but I know that you need a lot of ferritin molecules to form the thing. But all of those are also-- and again it's only 4 out of 52, but they're all proteins that exist in relatively high abundances. So could you also be making the argument that proteins that are more likely to have high concentrations, and therefore a higher probability of aggregating just because it's a prime molecular reaction could favor binding to groEL?

ELIZABETH   Yeah, I even thought about it in terms of they certainly are abundant. It could be, I just don't
NOLAN: know.

AUDIENCE: The experimental setup also biased it towards more abundant proteins.

ELIZABETH   Yeah, so could that have happened in the experimental setup? It's a possibility. So we learned
NOLAN: that what EFT was about 10% of all rhibosomal proteins. So that's something also to keep in mind, and a good thought there.

So what else can we learn? One more observation from these experiments before we move on to DNA K J. So recall last time when we talked about the actual pulse chase experiment, they took samples at multiple time points. And so why did they do that? You can imagine doing this analysis not just at 0 minutes and 10 minutes, but at a variety of time points and ask, if we compare gel to gel and we compare spot to spot-- so going back, these spots are labeled many of them in here-- we can ask the question, how does the intensity of that spot change over time? And what does that tell us about the interactions of that polypeptide with groEL?

So just for example, here. Example, just imagine at time equals 0 we see some protein or polypeptide x. So then what happens at, say, time equals 2 minutes? If we do not see it, let's

consider two options. Do not see x, maybe we conclude that x dissociates quickly or folds relatively easily.

Imagine if we do see x after 2 minutes here, maybe the conclusion is x is not yet folded here. And then we can imagine doing this at different time points, and they went out to 10 minutes here. So maybe if we see x at 10 minutes, the conclusion is x is difficult to fold.

And too, we want to think about these time points also from the standpoint of what we saw in terms of the residency time of a polypeptide in the groEL chamber. So we saw from the various models that that's somewhere on the order of 6 to 10 seconds. So there can be multiple binding and release events that occur.

So in this paper, what the authors did is trace the spots and compare the intensities of the spots over time. And you can do a little exercise from these gels looking at spots they circled and just ask qualitatively, what's happening to the spot? Is the intensity staying the same? Is it being reduced?

So for instance, it's easy to look at spot number 22 here at 0 minutes versus spot 22 at 10 minutes. And what do we see? Does it look the same, more intense, less intense? Less intense, right?

What about spot number 12 at 0 minutes versus 10 minutes? They look quite similar by eye. So you can imagine doing this type of exercise through each gel and actually doing it quantitatively using some instrumentation.

So what do they see? Effectively in this, they divided the data into three groups based on certain trends. And that's shown here where what we're looking at is the relative intensity change versus time. So you can imagine at some time point that spot has a maximum intensity that they've put at 100. So we see the three groups here. And the question is if we look at these as groups, what do the data show?

So in group one, we see examples where the spot at time equals 0 is at a maximum, and then the intensity of that, so spots decrease over time. And the other thing we see is that at some time that isn't very long, the intensities go to approximately 0. So we're not seeing these polypeptides bound any longer.

And then effectively what we want to ask is do these polypeptides have any similar features?

And what the authors observed is that the polypeptides falling into this group showing this behavior are smaller than 60 kilodaltons. And as shown here, they're seeing them completely released over the time course of this experiment, and in general within the first 2 minutes. So what does that correspond to? How they interpreted this was that these polypeptides are either binding groEL once or have several rounds of binding and ultimately reached their folded state in this relatively short time period.

So how does group two differ? Looking at these data, what do we see in group two that's different from group one? [INAUDIBLE] Yeah, we're seeing the relative intensity never go all the way to 0. So here we've gone to 0, here we see 20% to 30% as the cutoff.

So how are these data interpreted in this work, and what are the identities of these polypeptides? So similar to group one, these polypeptides are also all smaller than 60 kilodaltons. And how this behavior is interpreted is that even after 10 minutes, there's some fraction of these polypeptides that are still associated with groEL. So they haven't reached their native fold and are remaining bound.

What's going on in this group here, group three? This behavior is very different.

AUDIENCE: You see peak intensity is a little bit later than the rest of them, and they also don't go to 0 after 10 minutes.

ELIZABETH NOLAN: Yes. So these proteins are interacting with groEL because they were pulled down, but it looks like they're interacting at later time points. So we see this growth in terms of increase in intensity over time, and then they go down. And here we see 40% or higher. So they are not readily dissociating, binding at longer time points.

So one question here is are these dead end species? And within this work, the authors did some additional controls which there's some detail in the notes I'll post in lecture. But effectively asking, what happens if we add in groES, what happens if we add in ATP? Do we still see these species or not? And some of them were released under those conditions there.

So in summary, what we see from this is a method to look at chaperone substrate selection in the context of a cell. We see that groEL folds proteins over a range of sizes, but not really the small ones. So under 20 kilodaltons not so much, and over 60 kilodaltons not so much here, and that these polypeptide substrates have complex native folds.

So where we're going to close the chaperone unit is with looking at the machinery DNA K J.

And so we'll introduce that system and then look at a similar series of experiments where the substrate scope for this chaperone system was evaluated. So if we go back to the overview from the start where all of these players were introduced, this is where we are now.

So we're looking at DNA K and its co-chaperone DNA J. So these are downstream of trigger factor. What do we have for DNA K and J?

So these are heat shock proteins. DNA K is an HSP 70. So 70 kilodaltons, and HSP 70s are ubiquitous. So just to note, they're involved in a variety of protein quality control functions. So we have folding, as we'll talk about in the context of today's lecture in this module, but even rolls that range from protein transport to assisting with protein degradation occur. So here we have HSP 70 for DNA K and HSP 40 for DNA J.

So in this system, DNA K is the chaperone and DNA J is the co-chaperone, and DNA K is ATP dependent. So it's monomeric. So with this system we don't have a chamber like we have with groEL, groES, and it's ATP dependent. DNA J is the co-chaperone here.

So what happens in terms of this system? So effectively DNA J, the co-chaperone, scans hydrophobic surfaces of proteins or polypeptides, and it associates with them so it binds. And then what DNA J does is it delivers non-native polypeptides to DNA K.

And then how we think about DNA K is that DNA K binds and releases unfolded polypeptides. And this is another case where there can be multiple cycles of binding and release. So DNA K will bind to a polypeptide that has an unfolded region, there'll be some period of time that that complex exists, and then DNA K will release it.

And so in terms of where it likes to bind, these are typically six to nine amino acid segments that are hydrophobic. So it likes residues like leucine and isoleucine. And statistically, this type of region occurs about every 40 amino acids.

And for these segments, just to note that there's a range of binding affinities. You can imagine there's a variety of possibilities here. And what's found from studies is that the KD of DNA K for various polypeptides can range from about 5 nanomolar to about 5 micromolar, so by several orders of magnitude. In terms of size of polypeptide, it stated that DNA K has some preference for polypeptides on the order of 20 to 30 kilodaltons, but it can bind larger ones and it can bind polypeptides greater than 60 kilodaltons, as we'll see later.

So in this system there's another player that we need to think about, and that's this GrpE, or grip E. And what we have here is a nucleotide exchange factor. So any f, and it's also a thermal sensor. And what GrpE does is that it regulates DNA K binding to a substrate by inducing ADP release. So what we'll see is that the ATP and ADP bound forms of DNA K have different affinities for these polypeptide substrates.

So what we're going to do is look at the structures of the components of this system and then look at the cycle. And so if we consider DNA K, so we think of this protein as having two different domains. So there's an N terminal domain and a C terminal domain.

And in this end terminal domain what we have is the nucleotide binding domain, NBD. So this is where ATPase activity occurs, and this is about 44 kilodaltons here. There's a linker region, and then the C terminal, and we have the peptide binding or substrate binding domain. This is 27 kilodaltons.

So here if we think about this part just in cartoon form, what's observed is that there's a cleft for binding ATP or ADP. So ATP or ADP binds here, and this is also where the nucleotide exchange factor GrpE will interact, because that's its job as a nucleotide exchange factor is to help with that. So basically we have GrpE here. What we see in this domain, it's often described as being a beta sandwich plus an alpha helical latch. And the idea is that this latch closes in the presence of the polypeptide.

So effectively, if we look at this as a cartoon-- and we'll look at actual structures in a minute-- this peptide binding domain can either be in an open form, and this is the latch. You have the alpha helical part, here's the beta sandwich. And if there's some polypeptide to bind, what happens is that the latch closes and the polypeptide is bound here. So this is the closed form, and this pocket is hydrophobic. And that makes sense based on what we know about DNA K liking to bind hydrophobic stretches.

So let's look at some structures of DNA K. I present two slides of structures here, one from the assigned review and this other version. And I'll just focus on this one for here.

So here we're looking at the domain organization. What we have here is the nucleotide binding domain. So here's that clef for ATP binding. Here we're looking at the peptide binding domain. So the beta sandwich region is in green, the alpha helical latch is in yellow, and we see that there's a model polypeptide here, and this is in the closed form.

Here's another view of DNA K with a peptide bound. So we see the beta sandwich, here's the alpha helical latch. This depiction here from the review is showing the closed and open states, and closed and open is referring to the green area here. So don't get confused with the nucleotide binding domain and how these are shown. So what we see here, again, there's a bound polypeptide in this peptide binding domain, and here there's no bound polypeptide. And we see that now this alpha helical region is sticking up there.

So what about DNA J? You consider DNA J, we're just going to focus on the domain organization and just a more simplified view than what's on the slide. We have two domains for DNA K binding, and then for peptide binding.

So DNA K is going to go out there and find some polypeptide that needs the help of DNA J. It's going to bind that polypeptide and deliver it to DNA K. So effectively, it interacts both with the polypeptide substrate and it also acts with DNA K when delivering this polypeptide.

So just to point out DNA J is part of an HSP 40 family, and these are quite diverse. I just illustrate that from the range of different sizes, so from about 100 to about 2,000 amino acids. And all of these HSP 40s have what's called a J domain, and in this more detailed depiction here it's indicated these 70 amino acids at the N terminus are the J domain, and they're important for interacting with DNA K or another HSP 70.

So what about GrpE? This nucleotide exchange factor, so GrpE is a homodimer. And if we just look at one monomer, and then I'll show you the structure. So in '97 a crystal structure of GrpE with DNA K nucleotide binding domain was published, and this is what came from that. So just use your imagination, maybe I'll draw this a little differently.

Basically what we see with GrpE is that there's a beta sheet, and this is the C terminal region. And then what we see here is an extended alpha helix. And this is the end terminal region. And this is just a cartoon of the monomer.

So what happens is that the GrpE homodimer uses one of the beta sheets of one monomer to insert into that ATP binding clef here of DNA K. And when that happens, it forces it open there. So let's look at the structure, and this is something that actually puzzled me for quite some time, but there's been a recent update.

So this is a crystal structure of GrpE homodimer. So we see one monomer in blue and one monitoring green bound to an end terminal nucleotide binding domain of DNA K, which is

shown in pink. And so we see the beta sheet region of each monomer, we see the extended alpha helix. The C terminal end is here, the N terminal end is here. And I note that not shown in this structure is there is an unfolded region after the end here of GrpE.

And so we see this nucleotide binding domain interacting with one of the beta sheets. So a one to one stoichiometry. So the idea, as we'll see when we go forth with the cycle, is that GrpE is inserting the C terminal beta sheet into the nucleotide binding class of DNA K. And this happens for the ADP bound form, and it facilitates ADP release.

So what's going on down here? Why is there this extended alpha helix? And I'll just note there is a study just in the past year where interactions between DNA K and GrpE were studied in some more detail. So they used some biochemical experiments, some cryo electron microscopy. And what they learned is that the interactions between GrpE and DNA K are more complex than what's seen here, and what they observe in their cryo EM is evidence for this N terminal region interacting with the substrate or polypeptide binding domain of DNA K.

So there's some dynamics and flexibility that we can't appreciate from this crystal structure. And so that begs into question, how else is GrpE facilitating this cycle and modulating confirmation and function of DNA K? So you're not responsible for these details, but if it's something you're curious about I've included the reference. So effectively GrpE accelerates the release of ADP, and that in turn promotes binding of ATP.

So what is the functional cycle? And we'll look at this depiction here. There's another depiction in the notes from the reading. This is the current model. And in this model, we're going to start here.

So what do we see? We have DNA K in the ATP bound form. So we have the two domains-- the nucleotide binding domain, and here the polypeptide substrate binding domain. And in this cartoon, we see that alpha helical latch is open, so no polypeptides bound. And what we also see is that the ATPase activity here is very, very low. So DNA K is not hydrolyzing its ATP.

So then what happens? DNA K-- sorry, DNA J, the co-chaperone, has found some polypeptide substrate-- indicated by this S-- that needs the help of DNA K. So J binds the polypeptide substrate, and it delivers that polypeptide to DNA K. So what does this cartoon tell us? It tells us that J is interacting with K, and here we see the polypeptide substrate being delivered.

So when DNA K is in the ATP bound form, it binds peptides with relatively low affinity and in a

reversible manner. So there's fast exchange, that polypeptide's going to come on and off. And when DNA J binds and delivers the polypeptide, it activates the ATPase activity of DNA K.

So that's indicated here. So the ATPase activity is enhanced substantially so you can compare the values for some quantitative insight. There's ATP hydrolysis. ATP hydrolysis results in release of DNA J and PI.

So now what do we have? ATP's hydrolyzed, and now we have ADP bound in the nucleotide binding domain. And what do we see? The latch has closed-- open, closed. So like what we saw in the structures with those model polypeptides mound, we have the substrate clamped in this latch.

So here we have a form of DNA K that binds the polypeptide with high affinity and slow exchange. So this state is considered to be long lived, on the order of 10 to 15 seconds. So the question is, if this is binding the polypeptide with high affinity and slow exchange, how do we release it? And that's where this nucleotide exchange factor GrpE comes into play.

So here comes along GrpE. GrpE binds. GrpE binding results in release of ADP from the nucleotide binding domain.

So GrpE is inserting its beta sheet into that clef, and it looks like something else is happening with that long alpha helix to facilitate this. But this was drawn before that 2015 study, so we just see it interacting here. But imagine that this region here is maybe interacting down here and doing something to facilitate peptide release.

So now what? No nucleotides bound according to this model. Since the ADP is released, ATP binding is facilitated so ATP can bind. And what do we see? There's release of the peptide, release of GrpE, and this cycle can start over again.

So effectively, the release of ADP is accelerated about 5,000 fold from the action of GrpE. And so GrpE is called a thermo sensor and can begin to think about why that might be. If, say, there's condition of heat shock or stress, maybe the cell wants DNA K to be able to hold on to this polypeptide rather than release it. So GrpE won't be doing its job under those conditions. So another example of ATP binding and hydrolysis modulating activity of these chaperones.

So we need to think about what are the substrates for DNA K J, and what is the chaperone system doing? So we define possibilities as foldases-- like what we saw with groEL-- holdases, unfoldases, what's happening here? And so in thinking about the in vitro substrates, what are

the experiments we're going to do? Or sorry, in vivo substrates.

So can we take the method used for groEL, groES, and adapt it to this system? Are you convinced that method was useful, or are you down on that method?

AUDIENCE: It can probably be adapted.

ELIZABETH NOLAN: Yeah, right, it can be adapted. So can imagine again going to do a pulse chase here, and can imagine the same experiments where we have our E coli with no methionine to deplete. We can pulse with radio labeled methionine-- again, 15 seconds, 30 degrees Celsius-- to let us see newly synthesized polypeptides. And this gives us a way to ask what newly synthesized polypeptides did DNA K and J act on.

Then we can chase with excess unlabeled methionine for 10 minutes. And again, can take samples at varying times. Do rapid lysis. And in this case, rather than using EDTA to quench, what they did is do rapid ATP removal by adding an ATPase here. So just to realize that there's theme and variations in terms of how you can quench these.

So what do they find? And we'll go over the data in more detail starting on Monday. And what do they need to do to find that? So in this case, they need an antibody to DNA K if there's going to be an immuno precipitation, right?

So in these experiments, effectively we're to this point. They immunoprecipitated with their DNA K antibody. Of course, the specificity of this antibody needed to be studied, and then they used SDS-PAGE to analyze the immunoprecipitates.

And so what we'll see when we discuss the data next time, the experiments were analogous to what was done with groEL, groES, but a few differences. They were less sophisticated in terms of the approach. So they use just standard 1D STS page rather than 2D, and it didn't go through the process of doing trypsin [? digestion ?] mass spec to identify the polypeptide. So it's more of a qualitative look.

But we're going to ask starting on Monday what did they learn from analyzing these gels about the substrate scope of DNA K J? And then we have to ask the question, how does that help our understanding in terms of the type of chaperone activity that's occurring? So with that, I'll close. We'll end the chaperone unit with those experiments on Monday, and then we'll

transition into module 3, the proteasome and degradation chambers.