

**BARBARA IMPERIALI:** So what we're going to do today is we're going to finish up a bit of transcription, and then I'm going to talk largely about transcription control. Because it's tremendously important that we actually not just know how to transcribe, but we know how it's controlled, and how ultimately the appropriate messenger RNA is sent out for a translation into proteins. So transcription control is a very important component.

So let me just start with this first. So these are typical of questions. And they look like, gosh, I don't have enough information to answer this question. And it's basically about what's known as the transcription bubble, which is the portion of the double-stranded DNA, where transcription is occurring. And so a transiently, the double-stranded DNA is opened up for the RNA polymerase, which remember has inbuilt helicase activity to start transcribing the gene.

So let's just say we have the information we might give you is that transcription starts here at this point on the double-stranded DNA. And what we want you to know is that you have all the information you need to know which of the two strands is copied. Have you had a chance to think about that? Does anyone want to give me a good answer to that? Yes, here.

**AUDIENCE:** The bottom [INAUDIBLE].

**BARBARA IMPERIALI:** And why is that?

**IMPERIALI:**

**AUDIENCE:** Because it's [INAUDIBLE].

**BARBARA IMPERIALI:** Right. So that reading things 3 prime to 5 prime, and making things 5 prime to 3 prime, is all you need to know to answer this question. So you would straightaway know, OK, we're going to start somewhere here, but we're actually going to start on the lower strand, because that's the one 3 prime to 5 prime. And we're going to make the new transcript in the appropriate direction, and know what it's going to be. So you can only read the bottom strand in this case. But watch out, because you might be reading the top strand in the appropriate direction as well. But what we've told you already is where the start site is. So you're going to know that information.

If we'd put, for example, that the start site was over here, you might have different answers to this question. So the only singular piece of information you need is that you read 3 to 5, and

you make 5 to 3. Then you can also fill in the bases to know what the transcribed sequence looks like. And I always recommend when you're filling in bases, you just write 5 prime to 3 prime. So you really are following properly which is being read, and the direction it's being read. And so if you were asked what the new messenger RNA sequence would be, you'd get it right.

And here I just have a few questions that will help highlight the differences between the transcription process and the replication process. So can you guys just read each of these, and see whether they look like rules that apply to transcription, or they're things that are not true about transcription. So just give you one second to think about that, who? OK, who would like to give me an answer? Someone over in this part of the room, I haven't heard as much-- yes?

**AUDIENCE:** Makes a complete copy of one strata.

**BARBARA** DNA. OK, so the correct answer is that we only transcribe about 1.5% of the genomic DNA.

**IMPERIALI:** We're not transcribing the whole thing. We're only transcribing the bits that we need to make proteins. So the correct answer is C, because we don't make a copy of the whole of genomic DNA.

But these are all right. We have a different set of nucleotides, or rather a difference in one of the nucleotide triphosphate building blocks. Remember that RNA polymerase does not require a primer. That was a complication when we looked at replication, because we had to paste in primers, made often by the primase, RNA primase. All RNA polymerase is much cleverer than DNA polymerase, because it has the polymerase activity. It has the helicase activity. There isn't needed a topoisomerase, because we're only opening a little bit of the double-stranded to copy. I'll show you a movie in a second.

And it also has a 3 prime exonuclease, which means that RNA is able to do its own proofreading, just like the DNA polymerase. So the error rate is similar to the error rate with the exonuclease activity for. DNA so that's about 1 in 10 to the 5 to 10 to the sixth.

Now I'm going to talk in a minute. There's not the same cadre of repair enzymes that we have for DNA. And that's a curious thing, until you start thinking about what the reality is. So let me try to lead you to my thinking. So with genomic DNA, that's the copy of the DNA that's in the nucleus that has to stay good. So if there's any mistakes, we need them cleaned up. Because otherwise when we replicate a set of all of the double-stranded DNA, there will be an error in

the progeny, the daughter cells.

For transcribing RNA, an error rate of about 1 in 10 to the 5 to 1 in 10 to a 6, is just fine. That piece of RNA ends up, after a few of the processing steps that I'll describe to you, ends up leaving the nucleus to be made into proteins. And it's not such a bad thing if you have a little bit more error in that RNA. Why is that? Over here, or up there actually. Yeah?

**AUDIENCE:** It has a shorter life, so it's not going to mess up everything for the rest of its life.

**BARBARA IMPERIALI:** Right, so if you have a bit of an error, maybe you make a new protein, but it's not full length, or the protein you make isn't perfect; it's OK. Because there'll be other transcripts that are correct, and then after you've done making the protein, then you're just going to destroy that RNA, because it has a transient lifetime because of the structure of the RNA, and the nucleases that chew up the RNA. So this is an acceptable error rate for RNA polymerase. Remember, it's not an acceptable error rate to have in your genome. OK, it's just too big.

Any questions about that? Does that all make sense? OK.

All right, so now we want to talk about transcription control. But before I do that, I do want to very quickly show you something, because I feel like it really caps off the transcription part. And I'm going to show you only about a minute of this, because once again, I love the sound effects here. And I can't turn down the volume.

This is one of those animations showing transcription. And it basically shows on the double-stranded DNA, lot of things accumulating to make a decision with respect to starting. But now, the whole complex, the RNA polymerase is just screeching through the DNA, making that messenger RNA. And you're only unraveling just a little bit of DNA around that transcription bubble. So you wouldn't need topoisomerase in this case. You're only copying one of the two strands. We saw how to identify which. And thus, the new RNA strand is basically falling out of the complex. You can see [INAUDIBLE] it's moving a lot. So that gives you a feel for things.

When I hear that music, I'm just sort of waiting for the whole thing to crash somewhere. But who would know, right? All right.

So at the beginning of that animation a few things were highlighted, where there's the double-stranded DNA, and a collection of entities starts clustering around where the reading is going to take place. So this thin, black strand is the double-stranded DNA. So some of the entities

accumulate quite close to the start site. There is a section of DNA known as a TATA box that we spoke about, at the very end of the last lecture. But there are a number of transcription factors that are quite important.

But then you might also, if you go back and look at that animation, there are sections of DNA that are at quite a distance that also help really regulate and promote transcription. So many factors regulate where the transcription occurs. First of all, is there a promoter site right near where we need to start? That causes a fair amount of collecting of complex components. But then there may be other things at a distance. So the promoter region might be located quite near the transcription start site, but then there are also enhancers that can be located at quite a big distance away from the start site that also play a role.

Why do we need this much control? Because we don't need to be making the RNA for every protein all at the same time. Certain times in cell cycle you'll see, I only need to make one protein or a different protein. So we need all of this control to decide when transcription occurs, when do we need to make the messenger RNA to make our favorite new protein that needs to be made. And you'll learn a lot in signal transduction and cell cycle, where we really show you how a lot of the housekeeping genes are all fine. The proteins are all there. But at a certain stage, we need to make more of a particular protein. And that's when the transcription control comes into play.

And you'll commonly hear about things called transcription factors. And those may be the proteins, for example, that regulate that transcription should start. In other cases, there may be times when transcription is turned down. So we have things that activate. So we start transcription, we make that happen, or others that repress. So they turn down transcription. So it's all about when you need to start, which is controlled by external factors acting on the double-stranded DNA to send the transcription complex, making the new entity.

All right, does that make sense? We can't have a dysregulated system, or if we do, then we have problems with, for example, proliferation of cells.

All right, so what I want to talk about are the key things that regulate transcription, and the key things that we do to the messenger that's been made. So what you're seeing here on this slide is just where we are in the process. We've seen replication. We're now at the transcription step. But there are many steps to go in eukaryotes before that transcript can leave the nucleus. All right? And I want you to remember the difference between eukaryotic and

prokaryotic cells. This was just a picture we saw very early on.

Prokaryotic cells, like bacteria, do not have a nucleus. They have an area called a nucleoid, but they don't have a discrete membrane encased organelle, where the processes of replication and transcription occur. In contrast, in eukaryotes, there is a discrete area of the cell, the nucleus, that includes all the machinery for replication and transcription. And it also includes the machinery that takes a pre-messenger RNA into a mature messenger RNA that can leave the nucleus.

If you send that pre-messenger out there, it's going to have a lot of stuff wrong with it. It's not going to be ready to face the outside of the nucleus. It's going to be readily degraded. It's not going to have the full information. So I want to talk to you about the processes that are put in place for this conversion from the pre-messenger RNA to the messenger RNA, which we don't have to think about in prokaryotes, the small organisms without organelles. All right, so, and this is the foundation of transcriptional control.

All right. So, we've talked about promoters and enhancers. Those happen early. That's all about making the pre-messenger RNA. But now we have to discuss some aspects that are also critical for making the initial pre-messenger RNA, and that's chromatin remodelers, or chromatin remodeling. Because in order to transcribe anything, there's a lot of ground to cover with respect to unwrapping the chromatin, the chromosomes. Because they're all packed up in such a way that you can't possibly do any transcription there, because they're too tightly packed in order to be accessible for the transcription machinery to tackle it.

So I show you parts of that machinery up here. These are the nucleosomes that make up chromatin, which makes up the chromosomes. And in order for you to even be able to start transcribing, you have to unravel that part, those complex structures that are tightly wrapped up. You've got make the double-stranded DNA accessible, otherwise you can't break into it.

So there are two things that also contribute to allowing transcription, and those occur both at the DNA level, and the histone level. And I'm going to talk about the histone level changes first. I want you to recall that histones are proteins that have a lot of positive charge, by virtue of the fact that they include two of the positively charged amino acids, arginine and lysine. And if you're curious about those structures, you can go look back at the table of the amino acids and see that those guys are always positively charged.

The reason we use histones as the core of the nucleus home structure, is they're very

positively charged, and they neutralize the dense negative charge of the nucleic acid. Otherwise we couldn't pack it up as tightly as we do. So changes that occur at the histone level, and their remodeling of the chromatin in order to promote transcription is modification oftentimes to neutralize those charges.

So the one most obvious one I'll show you, and then there are others where you add methyl groups that dampen down the charge. But I'm just going to show you the very obvious one. So let's just look at lysine in a protein. That looks like this. That is a terrible drawing. It has a positive charge for bonds to nitrogen. In order to neutralize that charge, there are enzymes that acylate, or transfer an acyl group to turn this positively charged amine into a neutral amide.

Let me draw that side chain, because I think it makes much more sense to understand it, this particular thing from a chemical perspective. So we still have the n but it is now part of an amide. So the charge has been neutralized on the nitrogen. If the charge is neutralized on the positively charged residues in the histones, the DNA will be encouraged to unravel from those. It makes good chemical sense that that's a good way to start on packaging DNA.

The other modification occurs at the DNA level and it's methylation of cytosine. A little harder to explain from a chemical sense, but up there in the corner of this slide I've shown you the pictures. So methylation on cytosine would look like this. This encourages stabilization of the chromatin. So what would that do to transcription if I've stabilized the chromatin? It represses it. It turns it down.

So these are pairs of changes that act in opposite direction. So we'd go down-regulate transcription. OK, so DNA methylation causes the chromatin to be a bit more compact, more stable. It's much harder to unravel the DNA. It's obviously then harder to start transcription. In contrast, modification of the histone proteins to neutralize their charges destabilizes and up-regulates transcription, because it's allowing us to open up the nucleosomes in order to make the DNA available.

Does that all make sense? So we have two counter-balances that play in each direction. OK? All right, so that's chromatic remodeling.

The next two transformations I'll show you may certainly look fairly complicated. But I'm going to describe them to you, and show you what the changes are, and how they would contribute to stabilizing the transcript and finishing the transcript up, to make it ready to leave the

nucleus, to go out to the cytoplasm, where the machinery to translate proteins is. Because you want to remember the ribosomes that we're going to use on Friday aren't in the nucleus. They're in the cytoplasm.

So there is a variety of events. There is what's known as 5 prime capping. So that is going to be a change. Let's just say this goes down. This is the base. Remember 1, 2, 3 prime, 4 prime, 5 prime. It's something that's happening to this end of the messenger RNA that stabilizes it. So 5 prime capping is important for one end. And then at the other end there is polyadenylation of the 3 prime end.

OK, so let's just look at these one at a time. And let me convince you that they are important changes in the transcript that will preserve its identity, and in fact, give it a little bit more information. Because indeed, the 5 prime capping modification actually is a signal later on, when the transcript leaves the nucleus for protein translation. But in general, both of these changes mechanically protect the ends of the part of the gene that you're going to want to translate. They basically leave that piece of gene in the middle, where it's not going to be nibbled up, it's not going to be degraded. Because the biggest threat to the messenger RNA are things known as exonucleases. Exonucleases chew down nucleic acids from the two ends.

All right. So let's look first, and it's kind of wild and crazy chemistry, at the 5 prime capping. And then we'll discuss the other two. And then I will finish these types of changes with splicing, which is really cool and extremely important. But the transcript, the pre-messenger RNA has to go through all of these steps before it's ready for nuclear export. So in 5 prime capping, you have this strand of pre-messenger RNA. Everything looks pretty happy. It's all in one piece. But the first thing that happens is three phosphates are added to that 5 prime OH group. So that's the start of this process.

And this process actually happens while you're still transcribing the double-stranded DNA. It's actually already going on. As soon as that component of the newly transcribed pre-messenger RNA is made, things start happening at the 5 prime end to protect it. At that stage, then once those three phosphates are put on there, and nucleobase is added backwards. So there's a couple of functions. This all looks pretty strange. The rest of this still looks quite good. It looks fairly intact.

But then the next thing that happens is that the guanine that's here is-- you do not have to remember this stuff, I couldn't remember it. I just want to show you how weird and different the

5 prime end looks. It doesn't look like a strand of RNA. So the guanine is methylated. And then a couple of the riboses, those sugars that have an OH usually at 2 prime, get methylated. So we've created this entire thing at the 5 prime end, known as the 5 prime cap that looks nothing like regular messenger RNA. And that protects that end of the messenger, and makes it safe from a variety of insults.

So let's take a look. Why does it happen? So the first thing is it stops nuclease activity. Because the nuclease could look at it, and go, I don't recognize any of this happy mess. I'm not going to chop down this component. It's too foreign to me. So that's the primary thing. But then it's actually quite important for regulation of when that messenger needs to leave the nucleus. There are proteins involved in helping export the messenger RNA, when it's ready from the nucleus. So it's an important signal or recognition element. It marks this thing as a messenger RNA that's going to be important in translation. It gives it an identity. And then finally, it can actually in the next step promote translation.

So all of these things are useful. So we do a number of these unusual transformations, but they all have a reason, and they all carried out in the nucleus purposed to protect the 5 prime cap of the messenger, and actually make it ready for its next tasks. All right?

The next thing that happens is what's known as polyadenylation. So here's it. This would be the 5 prime cap. In the middle would be your gene that you're going to transcribe. But at the other end there is an enzyme that puts on a lot of adenine nucleotides, and basically adds to the other end a lot of A's. And when I say a lot, it can be hundreds. And it just promiscuously keeps on adenylating.

And now what's this bit for? Once again, it protects from exonuclease activity from the other end of the strand. Because if you have exonuclease activity, you might start munching away at these As. But these aren't the important parts of the messenger, right? These are just additions. It's kind of a buffer. It's like something to do while you're waiting. Oh, I'm going to chew up some As. But you're not messing up the messenger RNA that you need at the end. So it's adding some dummy sequence that will be handy there.

It contributes to stability. The tail is shortened over time, but it's non-coding, so that's OK. And actually though, when the tail is short enough, the polyadenine tail kind of acts as a bit of a timer. Because once the tail gets short, you start chewing into the transcript, and you basically end up with a degradation of the transcript. But it gives you time in the cytoplasm for the gene



to be translated into protein. Does that all make sense? So it's basically like an egg timer, just watching, watching; back, back, back. OK, this transcript has been out here long enough. We've made all the protein we need. Now we're going to chew up the coding part. And then finally it's actually a good marker again for leaving the nucleus. And then finally it's kind of a cool tool for technology, because there's lots of interest in characterizing what's known as the transcriptome.

OK, it's all well and good to characterize a genome. But it's a lot of work, 3.2 billion base pairs. And only a part of it, a tiny part of it, are parts of the genes that encode the proteins. So what you'd really want to know, if you want to analyze the genes that are going to become the proteins and maybe look for defects, is look at the things that are going to be transcribed. So you can use the fact that there is this poly-A tail on the transcripts that are going to leave the nucleus with something that will pull them out of the mix.

What would I use? Let's say I've got a resin bead, polystyrene or some favorable polymer. And I can attach covalently nucleic acids to that bead. What would I add there to fish this lot out? Yeah, up here? Yeah.

**AUDIENCE:** [INAUDIBLE]

**BARBARA IMPERIALI:** Yeah, a lot of Ts. So I'm going to make this. I'm going to put on a bunch of Ts, and I can do this a lot. This is done very, very commonly. And then I'm just going to fish out everything that is part of the transcriptome, not the genome. So I've got a much smaller job then to find out errors in the genes that encode the proteins, than if I was going to start with the entire genome. So pretty cool tool, and later on we're going to see how this can be used. It's actually used in concert with an interesting enzyme that comes from viruses called reverse transcriptase. But that's a story for a later day. But it is a cool story.

Now the last thing that we do in the nucleus is arguably the most important. And that is splicing. OK, all right. Once again, this is a picture with a lot of moving parts. But I want to convey to you the point of splicing, as opposed to you knowing all the little details. It was noticed for a long time that the messenger didn't always correspond to the original transcript. And this is the case eukaryotes, not in bacteria. But that there was quite a lot of processing done, not just to cap the ends, but by cutting out chunks of the transcript, so removal of segments of the transcript.

And a lot of the seminal work was done by Phillip Sharp, who is a member of our faculty in the

Biology Department. So we're very proud of this. This was actually the topic of a Nobel Prize in the '90s. And so what was noticed is that if you had a gene, and let's just put it here, 5 prime to 3 prime. And I'm going to name these. And then I'll explain to you what they are.

These would be called exons. These would be called introns. And this is another exon. Outside, inside; that's the way to remember them. And what happens in splicing is that the introns are chopped out. OK, so your gene then ends up being, if this is exon 1 and exon 2, not represented by this entire thing, but eliminating that middle component. And remember all the time we have the cap at 5, we have the poly-A tail ad infinitum, the 3 prime end. So we've just changed the middle.

And it was recognized through bioinformatics analysis that basically noticed the pieces that was chopped out and the pieces that ended up being joined. And protein splicing occurs as a sequence of events that ends you up with the protein being spliced together through a series of rearrangements that occur on the structure of the messenger RNA. So there is an internal rearrangement. A new phosphate is made. And then there's another rearrangement, where there's a new bond made between one end of the pre-messenger RNA, and the other end of the pre-messenger RNA, to give you the exon rejoined, ready to be read in translation.

Now why is this so important? So I want to go here to a few numbers that I think are very pertinent, which I had to drill my colleagues for, because I didn't know all the numbers. So when the genome was sequenced, we were a bit stunned, because the number of coding things that are genes that are going to be turned into proteins was much smaller than we anticipated. For man, it's currently at about 20,000 genes. That's 20,000 mature transcripts that could make proteins. Fly? Not much smaller, arguably they're pretty smart, 16,000. Yeast 6,000, a bacteria at 4,000. There doesn't seem to be enough of a difference.

But a huge amount of diversity is introduced into the transcripts by different splicing events. Because if you make one transcript, that would count as one gene. But if there is a lot of opportunities for difference splicing activity, you can piece together new transcripts that will encode proteins differently. And I think this next slide will show you an example of a transcript that has several different introns and exons.

So you can see here across, there's a blue, green, red, another blue, and an orange exon. But depending on which pieces are spliced out, we'll have different proteins end up. So it could be a way, for example, to think of a practical use, to have a protein that is either secreted as a

soluble protein, or left in the membrane with a membrane association domain, or left in the cytoplasm because it has no way to be secreted. So it could be a way to make three proteins that are in completely different places of a cell, or have different functions.

It's also very important in tissues, because we splice proteins differently in muscle, or liver, or heart to achieve different outcomes in the proteins that we make. So basically, it's a source of huge diversity, and it gives us the genetic diversity that means this 20,000 can be a much bigger number. And then with the post-translational modifications I talk about when we come back later on, you'll see we've got lots of ways to diversify that 20,000. That means these aren't literal numbers. They're quite varied. But what you want to remember is *E. coli* doesn't do splicing. There's no opportunity there. It's much more limited in yeast, so is the post-translational modification. So these numbers settle out quite differently from what they look like by looking directly at those numbers.

And I wanted to give you an example. You don't have this in your slides, but I was thinking about it the other day. A colleague of mine, Professor Pentelute in chemistry, works on trying to reprogram genes to overcome a disease known as Duchenne muscular dystrophy. It's another genetic disorder. It's X-linked. So it's much more serious in males than in females. Because if you only have one bad copy of the gene, you can do OK with this disease. Whereas in the male, you'd only have the component of the gene that's on the X chromosome. And it's a defect in RNA splicing, so directly there.

So this is a biopsy of muscle where red would be good muscle cells, whereas the white would be fat cells. And they actually weaken the integrity of the muscle, because the muscle doesn't develop to have all the well-defined muscle cells that are important for muscle tensile strength, and contractility, and everything else that muscles do. And it's all related to a protein known as dystrophin, which is a huge protein. And it's a critical structural protein that's important to maintain the cellular membranes within the muscle cells. So if dystrophin is no good, the cell membrane integrity is no good. And you just end up with losing the muscle cells, a cost of replacing them with fat cells in the muscle.

And here's a really amazing number. The gene that encodes dystrophin has 79 exons. So you can picture the opportunities for things to go wrong, and it's actually a defect in the place where the splicing happens. So a splicing event cannot happen. And so the protein at the end of the day is not good. So there's a lot of efforts going on, antisense efforts, and even other gene therapy efforts, and also nowadays there's obviously a large focus on CRISPR-based

gene editing. But remember this is a serious debilitating disease of all muscle tissue. It starts to be noticed when the babies are toddlers, because that's really when they start to engage muscle strength. So around the age of four it's actually noticed. And then the life expectancy is in the 30 to 40-years-old range. But it's also a terrible lifestyle, because the lungs don't work. So many things rely on muscle strength. OK, good.

All right, so I believe that's the end of that. But I want to give you an introduction to translation, so you'll see what we have in store for Friday. And just get you back to this picture, we've done everything we need. We've done all the processing. We're here, and then the mature messenger RNA can finally leave the nucleus to be ready for translation. And the cap-- there is a complex that binds the 5 prime cap that's tasked with helping export that transcript through the nuclear pore, which is a fairly large opening, outside into the cytoplasm.

And I want to show you one other thing and I hope I can get this reliably, because with this I'm going to actually show you some of the moving parts. There's a little thing on the sidebar now, called a short translation. And so what you're seeing here, the black line is the messenger RNA. OK, so that's the thing that's finally out there in the cytoplasm ready. The pale green and yellow are the two components of the ribosome, which is a soluble organelle that assembles on the double-stranded DNA based on a few cues.

The dark blue components are transfer RNAs that bring in new amino acids. And what you see threading out here is a new protein strand that is being translated. This particular translation is to make a protein that becomes membrane bound and secreted. So I'm not going to go further than that. But there's a lot of those light blue proteins that are actually helping escort the transfer RNAs to where they're needed to continue building the proteins.

So we've seen the messenger. We'll learn more about the ribosome, and will learn about the transfer RNAs. And you can also watch this to your heart's content. I just find it just a useful simple cartoon that tells us a lot. So what I really want to encourage you to do is read this section 14.5, because the next parts will make a great deal more sense.

So what I want to do though now is first of all, start with describing the players in translation, and just the way we did for replication, we're going to systematically ascribe the importance to each of these players. So what you see, there are molecular players and there are key steps. So the molecular players are shown here. So far, we've focused very strongly on the messenger RNA. You know all about that. You know where it's from. You know it's stable in

the cytoplasm for a while. You know it's got signals that tell the complexes where to go down.

Then we need transfer RNAs and amino acids, and we need the ribosomes. And ribosomes are made of a composite of nucleic acid and protein wrapped together. There's quite a big difference in prokaryotes and eukaryotic ribosomes, and I'll mention that when we get to it. But I want to focus for a few moments, first of all, on the landmarks that got us to finally in 2009 the structure of the ribosome. It's a cool development for everyone. But also where we came from in the '60s, when Crick and Brenner finally cracked the genetic code, and found out that three bases on that messenger RNA will encode each new amino acid that gets put into your protein sequence.

So I want to introduce you to the transfer RNAs that are really the most important element to start thinking about. And you can call these guys the decoders, because what the transfer RNA does is it can carry an amino acid on one end of its nucleic acid. But on one of the other loops within the transfer RNA is what's known as an anticodon that recognizes a codon in the messenger RNA, and basically prescribes what amino acids get put in.

So it was always a wonder how you went from the nucleic acid world to the protein world. The answer is you use a nucleic acid that you can load with an amino acid, but it can also recognize the messenger RNA that codes for the protein sequence. And that's why I like to think of them as decoders. But what this picture of the RNA shows you, is the beautiful structure of RNA, where at one end is where an amino acid would be attached. There's this cool kind of cloverleaf structure. And then at the other end is the anticodon loop that will actually recognize your messenger RNA.

So I'm going to stop now. But I really want to encourage you, just skim through that small section. It'll make a lot more sense on Friday. I've been waiting to talk to you about protein translation since we started this class. But it'll make a lot more sense. And there's some super cool initiatives in chemical biology now, where people have been able to completely hijack protein translation, and not put in just 20 regular boring amino acids, but actually put in all kinds of other amino acids. So we understand the system well enough to manipulate it. And this really harkens to the Feynman quotes, "If you can build it, you can understand it." That's the level to which we understand translation nowadays. That's it for today.