BARBARA IMPERIALI: OK, we'll get started. I've got everything turned on here. So, a couple of things. I've been mentioning that it's really kind of cool, I think, at this stage, where you've gathered enough steam in some of the topics that many things that come out in the news might start looking of interest or relevant to you. And I found this.

It had a news brief in *The Scientist,* and then I went to the original paper, which is in *Nature Biotechnology.* And if you recall, towards the end of the work on proteins, we were talking about phenylketonuria, which is a genetically linked disorder, where people cannot metabolize phenylalanine to keep the levels of fell phenylalanine in check. So what happens is the phenylalanine gets converted to a toxic material, and it causes a lot of fundamental physiologic disorders that are a lot of neurologic disorders.

So what a small company in the area called-- what's the name of the company? I can see it here. Anyway. Synlogic is a synthetic biology company, which basically engineers bacterial strains as probiotics that can be used to mitigate some genetic disorders. And so they've created a bacterial strain that can metabolize phenylalanine, so did a good job of sort of not letting your phenylalanine levels get too high. So in the GI system, this bacterium basically works on phenylalanine to metabolize it so that, then, people don't have to lead such a strict dietary regime and have way less risk.

So I think it's a really cool thing. This probiotic is in one, two clinical trials. It's being fast tracked. And the company, in general, is a synthetic biology company working on solutions to certain types of diseases that could either save you from having a lot of restrictions on your lifestyle or, alternatively, save you from taking medications and stuff. So that's something that caught my eye.

I want to remind you that I have now put the link to *The Scientist* in the sidebar of the website so it's much easier for you to grab it and take a look at what's in the news. There's stuff in the news every two or three days. And there are things that I think you'll find interesting that really relate to technology, engineering, and fundamental science that are related to biology.

The other thing I want to do is remind you that towards the end of the class-- but no time like the present, because it is sort of the equivalent of one of the problem sets but, in fact, worth a little bit more than the problem set. I want to encourage you to keep an eye on *The Scientist*

link and maybe pick out a topic that two or three of you would like to write a news brief on with a paragraph of writing on introduction and what the technology is, how technology has addressed a particular scientific or biological problem. And then there should also be a graphic that describes it.

Not stuff snipped out of whatever you're reading. Something that you create as a team to sort of describe a concept to people. There's alternatives with that assignment. You can also pick an interesting protein from the PDB, make a 3D print of it, learn how to print it, and go to the 3D printers in the maker labs, and print the protein, and then write a brief summary of what it is. And one other idea I had for the engineers' view is I would love a better, less clunky topoisomerase demonstration.

In particular, I'd really like one where you can snip the pieces apart, let the untangling happen, and put them back together. So some of you could work with a couple of colleagues and make something, which I know a lot of you are really keen on, which is why you're engineers.

OK, any questions about any of this? I'm trying to make sure that this doesn't creep up on you. It's just something you do. It's great to get awareness of technology in the life sciences because of how much contribution it makes. So if you keep an eye on things, you won't be forced to suddenly find something good at the last minute. You'll just have found something and go, that's the perfect thing to describe. All right? Questions?

And the other thing is if you're unsure, you can always let us know what you have chosen or what you think you're going to choose in the chat with us, and we'll say, yeah, looks like a good idea. OK, so let's move forward now. All right, so what I want to do, first of all, is just remind you-- it kind of flew by a little bit-- that DNA replication is bidirectional.

So what that means is wherever you have an origin of replication, you can replicate in two directions. And I was sort of falling asleep thinking about this. I sure wonder what happens at the other side of the circular plasmid when the machineries kind of collide and you spit out a brand-new copy of a circular chunk of DNA. But don't think about that too much. It'll probably keep you awake too much.

So this is circular DNA. So what organisms have circular DNA?

AUDIENCE: Prokaryotes.

BARBARA Prokaryotes. Remember, the eukaryotic DNA is linear. And actually, we're going to talk about a

IMPERIALI: conundrum with the eukaryotic DNA because of the ends of the DNA, the ends of the chromosomes and their copying, when we talk about telomerase. But this is typical of a bacterial circular DNA. It's usually super coiled and becomes uncoiled in order to be replicated.

So, obviously, going bidirectionally gives you twice the speed, because your roaring around the same time. The helicase opens up the DNA. DNA polymerase does its job. The pink strand here would be the leading strand. And there's obviously going to be one on both strands of DNA, which will replicate very well. But don't forget, you're going to have to deal with the lagging strands in both cases.

So in both cases, what you would do is stick down a primer in order to be able to build those lagging strands. Otherwise, DNA polymerase can't get a grip on the double-stranded DNA in order to carry out the synthesis. So in those cases, there would be a primer to set up the lagging strand, and then DNA synthesis can occur. So let's put that in.

Once there's a primer here, DNA synthesis can occur. And then what happens with respect to the pieces that we're dealing with? What happens with the primer? What do we have to do to get to a nice complete, intact strand of DNA? Which enzymes are involved? Yeah?

AUDIENCE: Ligase.

BARBARA IMPERIALI: Here, there would be ligase activity. We need that. What about at the other end? What do we do with the primer, and then how can we move forward? Someone else? What happens with the primer? It's oftentimes an RNA primer. You'll see in a moment that RNA doesn't need a primase, so it's very easy for RNA polymerase to stitch in little pieces. What do we have to do, though, to get an intact strand of DNA? OK.

AUDIENCE: Have to remove the [? RNA. ?]

BARBARA IMPERIALI: Yup. So you're going to remove the RNA. Then the polymerase, later on, when we keep going, can sort of build this piece. And then we'll have to ligate that, as well. So you want to remember all the functions of those enzymes that are involved in replication. It's a little worrisome that people-- I know you don't want to talk or you think that's an obvious answer, but it's really important that you have them at your fingertips, some of these enzymes that are involved in this process. Because they should start to become second nature.

When you have to make a full DNA copy of an entire genome, there's a lot of moving parts.

But if you start walking through the logic of them, they make sense. If I'm going to unpeel DNA, I need a helicase. If I'm going to keep it single stranded, I need single strand binding proteins. If I'm going to move forward, sure, I need the polymerase, but what does the polymerase need? It needs a primer in order to have double strand, because DNA polymerase only wants to lock onto a double strand to go start doing its job.

These complications with the lagging strands that are really annoying, but it's pretty remarkable nature has addressed this and is able, remember, to replicate DNA and bacteria at a speed of 100 base pairs per second. So that's what's going on, this entire process. Excuse me. 1,000 base pairs a second. I just want you to remember this process is slower in eukaryotes. It's about 30 to 50 base pairs per second. Obviously, when you're speeding, you make more mistakes, so there are more mistakes in bacterial genome replication.

Why does it not matter so much if there's a mistake in a bacterial genome? What do you know about bacteria and their lifestyles? Do they stick around a long time? No. So they divide quickly. They live and die quickly. So you're not having to keep an intact genome without mistakes in it for a long time, because you're just turning over bacteria. If there's a mistake, it probably dies out. Or heaven forbid, there's resistance to drugs developed.

And we'll talk about those later, because those occur due to mistakes and in bacterial replication. But in a eukaryotic genome, we have to preserve the integrity of the genome. So I'm going to talk about two things that are related to the accuracy of replication now, because that's a really important component.

All right, so the first thing is to think about, what's the basal rate of making mistakes of DNA polymerase? So for that purpose, I'm just going to put down a piece of DNA with its partner that's being synthesized, five prime to three prime. And I'll put in some bases, so A. So that would have had a T put in opposite. G, that would have had a C.

So these have two hydrogen bonds. These have three hydrogen bonds. And let's say we now have a C here. So we want to put in a G at the position opposite. DNA polymerase wants to add the next base pair. It should be a G, because it's going in opposite a C. It's being grown in the right direction, five prime to three prime. So the basal error rate is about 1,000 to 1. So 999 times out of 1,000, the right base gets put in. 1 time out of 1,000, you might put in the wrong rate.

So the error rate is about 1 in 10 to the 3. That's really all that's-- all that's at play here is

energetics, just how favorable putting in the right base is. But there's a slight chance that the wrong base will just go in. Because the energetics are sufficiently different, but you're going to make mistakes, just because of the thermodynamic balance. If you're putting in 999, you're going to get it wrong some of the time, just statistically, because the difference in energy between putting in the right base or putting in a wrong base.

So that error rate is too high. If we replicated our genome, 32 billion base pairs, and we had a 1 in 10 to the 3 error rate, we'd have a lot of mistakes in the genome, right? And we cannot tolerate that, because all those mistakes in our genome will then propagate to mistakes in our proteome, if we're in the right segments of the genome. So this is pretty unacceptable with respect to an error rate.

So one way that nature deals with this is that DNA polymerase actually does some proofreading, all right? So it has a proofreading function. So what do you do? When you're proofreading, you take a quick look at what you've just written and say, oh, yeah, that looks good. That looks good. So what DNA polymerase is-- it more or less reaches back to the base it puts in and checks that it's OK.

It can only proofread one base back. It can't proofread from work that's been done a long time ago. It can just proofread very recent work. And if it looks like it's the wrong base, DNA polymerase has an opposite function. It has what's known as a three prime exonuclease activity. So I'll write that down, and then we'll talk about what that means. Three prime exonuclease.

So let's say we put in, instead of G, we put in a T. That's bad news. So what it can do is it can reach back and cut off, from the three prime end, a single nucleotide, the one that just got put in, all right? And then allow the process to reoccur to get the right base pair in.

So a lot of enzymes will catalyze both forwards and backwards reactions. DNA polymerase, the energetics of such are that it is able to catalyze both the addition of a nucleotide and the removal of a single nucleotide, only if it's at the three prime end. Only if it's at an open end, where it's just been put in. So, remember, the DNA polymerase is still here, because its plan is to move forward and keep on putting in nucleotides. But it actually checks back.

You could picture DNA pol just sort of quickly looking over its shoulder at the work it's just done and realizing that's the wrong one. So what this does is brings the error rates down considerably. Goes in 1 in 10 to the 5. So that's way better. 1 in 100,000 is much better, and

that's pretty acceptable. So it means you're really making very minimal mistakes in the replication.

So this part, the proofreading, brings the error rate from 1 in 10 to the 3 to 1 in 10 to the 5. But it can only work during DNA polymerase activity, and it can only work on the most recent nucleotide that has been put in. So this is basically a summary of-- yes, question?

AUDIENCE: So are [INAUDIBLE] in prokaryotes and eukaryotes?

BARBARA IMPERIALI: They are, and they are similar-- actually, there's slightly less error rates in eukaryotes, because the speed is slower. So the opportunity to fix things is going to be a little bit better. So, in the end, your goal, really, is to bring your error rate between 10 to the 5 and 10 to the 6. But for bacteria, because the speed is so much larger, this is sort of the limit of it. But in eukaryote, it can be a little bit better, because the speed is slower.

So you could imagine, if you're quickly proofreading, you're doing a less good job than if you're slowly proofreading. But the enzymes that I'll talk to you about in a second in eukarya are to fix entire work that's already been done. We'll see that in a moment. And that's what really cleans up. These enzymes are called the guardians of the genome, and they are much more sophisticated in eukarya. So that's a very good point.

All right, so here's the general scheme. There is an extension. There's a mistake. So there's proofreading. The mistake gets taken out. And then you keep on extending again, all right? So this now becomes pretty good. But what we need to talk about now is, what are the enzymes that go to work-- and I'm leaving what's on that board there, because I'm going to come back to it-- to actually correct mistakes in DNA.

So let's talk about the guardians of the genome. Because remember that your DNA is your permanent record of what needs to be made. All right, so the types of mistakes that can be fixed with proofreading are only recent mistakes. The types of enzymes I'm going to talk to you about mistakes that are found globally within double-stranded DNA.

And these are mistakes wherever. They may be mistakes that didn't get corrected by proofreading. But more importantly, they're mistakes that occur due to some kind of damage on the DNA. So what kind of things might impact the integrity of our DNA on a day-to-day basis? Yeah?

AUDIENCE: Sunlight.

BARBARA IMPERIALI: Sunlight. So terrible stuff. UV light will actually cause some cross-links, and I'll show you one of those, that are very serious in DNA. What else might hurt the genome? So sunlight. People say, don't go out in the sunlight. They're right about that. What else? Yeah?

AUDIENCE: Radiation.

BARBARA IMPERIALI: Radiation is another form. So that's like radioactivity. Radiation is important, right? So if our ozone layer gets thin, there's more risk there, as well. What about barbecue? Yeah?

AUDIENCE: Harmful chemicals.

BARBARA IMPERIALI: Yeah. Awful chemicals, terrible chemicals, right. So chemicals. So these are things like polyaromatic structures that actually slip into your genome and cause mistakes in the reading. Or they actually physically modify the bases to make it a base that doesn't look like a base anymore. So these are commonly very reactive chemicals. So these are all very serious things to the genome. And so the enzymes that mitigate the damage to the DNA basically screen along the double-stranded DNA to look for defects.

Because if you have perfectly paired DNA, then you're not-- you're going to have a very regular structure. Whereas, if something has happened to the DNA, there's something wrong with the base-- it's not base pairing well or something has actually happened between bases where they're causing a bulge in the genome-- then these enzymes will come into play. About a few years ago-- actually, it was on a class day, so I always enjoy these.

The Nobel Prize was awarded in 2015 for the researchers who deciphered the mechanisms for correcting the genome through DNA repair mechanisms. So there are two basic mechanisms that I'll talk to you about. One is base excision repair. And the other one is a lot more serious. It's actually the entire nucleotide excision repair. So one fixes just the base that's gone wrong, but the other one takes much more of the structure out to fix it. So it's nucleotide excision repair.

So BER and NER, and we're going to talk about both of those mechanisms, because they're very fascinating. And they kind of lean on some of what you've learned already. So base excision repair occurs when there's a defect in the base. Maybe it's the wrong base. Maybe it's just been modified a little bit. So what happens in base excision repair is that the base gets-- all right, I'm going to just use the pointer, because my little spotlight-- it'll pop back. It's a bit

magical.

So base excision repair, only the base, a single localized base, is damaged. There is some chemistry, for example, that will convert cytosine to uracil called a deamination mechanism. And in fact, if you replace the cytosine with a uracil, then you get in trouble with respect to its base pairing with its appropriate purine partner. So in base excision repair, what will happen is that base will be detected. It will be flipped out from the context of double-stranded DNA.

If it's tucked in the double-stranded DNA, you can't quite cut the bond that's attached from the ribose sugar to the base. So the base gets flipped out of the DNA structure. And there's an enzyme known as a glycosylase that cuts the bond between the ribose and the base and gets rid of it. And then what happens is the rest of the nucleotide, just that one nucleotide, gets removed. And then DNA polymerase fills the gap, and the strand is sealed by a ligase.

So a glycosylase cuts the base out. There's a couple of enzymes that actually cut the ribose, the phosphodiester linkages out. And then the two enzymes, remember, that are important when we're making DNA, the polymerase and the ligase, work together to put a base into this position. And then they put the base back in, and then the ligase joins the gap.

So DNA polymerase will make one of the bonds. The ligase will make the other bond, all right? So that's base excision repair, and it's monitored by finding that there is a lack of integrity in the double-stranded DNA. It's only a base that's affected, so that base gets removed. The rest of the nucleotide is removed, but only one of them. And then they're replaced through the concerted action of polymerase and ligase.

Now, there is another mechanism that's much more serious, and this is very typical of the types of damage that get formed from sunlight and UV radiation, is when two thymidines are adjacent to each other, they will undergo, quite commonly, a chemical reaction where they form a dimeric structure. So there's much more wrong with the DNA strand in that situation. So those get noticed.

This thing's driving me nuts. Those get noticed as a real defect in the DNA. Base excision repair won't work. Why wouldn't it work? Yeah?

AUDIENCE: Because the two are bound together.

BARBARA Yeah, they're bound together. You can't peel them back out. You can't break in that structure.

IMPERIALI: So you've got to move to a much more sort of major fixing of the DNA strand. So there is a genetically linked disorder where the enzymes involved in nucleotide excision repair do not work. And when a child has a copy of both defective genes from the parents, both one from the mother and the father, it's impossible for them to fix these defects in the DNA. And they get a lot of physiologic problems like sort of scarring and sunburns, barely anything at all.

So this is a group of children that are so afflicted with this genetic disorders that they cannot go out in the daytime at all. So you'll sometimes see they're called Children of the Night, because basically, they have to flip their schedules. They just can't go near sunlight. And if they go out, they basically have to be covered from head to toe. And that's including their eyes, because you can get sunburn of the eyes. You can sort of see, in some of these pictures, that it looks really serious defects.

And this is just the external manifestation. The internal manifestation would be cases of skin cancer very, very readily. So if you don't have at least one good copy of the enzyme that does nucleotide excision repair, then you're in trouble. And the defects are called-- it's called xeroderma pigmentosum. And there's actually a lot of family groups that get together, because the best way is just to form a sort of social network so the children understand what each other-- the limitations that they all have. And they can play together and be on these sort of flipped schedules in order to avoid any sunlight.

So, in this case, it's essential to basically clip out a large chunk of the DNA. So what happens in this case is that the DNA is recognized, and then a large portion of it-- about a dozen nucleotides-- are cleaved out. And then, once again, DNA polymerase fills the gap, and DNA ligase seals the last gap. So DNA polymerase will be able to fill going from the five prime to three prime. There'll still be one gap, and then the ligase fixes it.

So I think these kinds of things that are done to mitigate damage on the genome are very important to understand, because this is happening all the time. Any minor things that get fixed, that need fixing due to sunlight, radiation, chemicals will be fixed through these methods to keep that rate, that error rate, in your genome down to, like, one in a billion or something like that.

All right, I'm just going to flash this up. I'm going to, very quickly. I'm going to give you guys a copy of this. But these are the components that you want to be able to understand the function of when thinking about DNA replication. So you don't have that on your slides, but we're going

to give you a copy so that you can really make sure that you understand all the moving parts and how they come together for replication.

Now I want to talk about one last conundrum with DNA, and that is the issue of telomerase. And this is particularly critical in eukaryotes that have linear chromosomes. Now, if you think about DNA being replicated, when you look at the two strands of DNA, as you approach the very end of the chromosome, you'll do just fine making the copy that's built five primes-- sorry about this.

You'll do just fine making the copy that is built in this direction, because it's the leading strand. And it's built five prime. Am I going wrong here? It's built five prime to three prime. Can someone help me out here? I'm losing my mind. So this piece is built five prime to three prime, so therefore, this was five prime and three prime.

But then on the other strand, you have a problem, because you need to put in a primer here in order to build the other strand. Does that make sense? So we've got to have put in that short primer, because DNA polymerase is will not work otherwise. And then we need to build this strand of DNA. All right, so what's the problem here with respect to these ends? What's going to happen next? If it's an RNA primer, what happens next?

We nibble it up, right, with the full intention of replacing that bit of DNA. But then what can DNA polymerase do? It can't do anything, right? DNA polymerase needs double-stranded DNA to hold on to so it could fill this gap. So what happens is every time you replicate DNA, you end up with a gap, with a small amount you don't quite copy. Is everyone following me?

And that's a problem, right? Because doesn't it mean every time my cells divide, my genes get a little bit shorter and a little bit shorter and a little bit shorter? So there are things in place that help. One important feature is that, usually, you don't have important genetic material at the ends of your genes-- at the ends of chromosomes, rather. There's sort of extra DNA that doesn't need to be copied.

But the basic theory, the whole theory about telomerase, is that for certain types of cells-- these ascend stem cells and germ cells-- there is an enzyme that can fill this gap. It's called telomerase. So in those cells-- what's special about these types of cells? We need to keep them good. They're what defines your starting DNA. The stem cells and the germ cells, like in the sperm and egg, have to have a good copy of DNA. They can't be getting shorter and shorter every time a new generation is born.

But once your cells are the somatic cells, the DNA gets shorter and shorter, because those cells don't have telomerase. And this is associated with theories of aging. So the cells you get, the ones that are finally the ones in your body, every time they divide, the ends of the chromosomes will get a little shorter. But there's no mechanism to replace those. And so it's associated with the belief that, at a certain stage, you've divided the cells enough times. And then you can't-- you're actually starting to nibble into important coding DNA.

Does everybody understand what would be the significance of that? Does that make sense? Yeah?

AUDIENCE: So once the [INAUDIBLE] are gone, is this [INAUDIBLE] able to divide?

BARBARA IMPERIALI: No. So the telomerase keeps the DNA in those types of cells in good condition. In the cells that divide daily, and your body wants-- the somatic cells. You will keep on shortening, but they'll just be more mistakes, basically. And those are the sorts of things that would be associated with an organism that's growing to a certain age, because it's just a certain number of cells. So the cells may not divide, or there may be mistakes in certain parts of the coding genome.

Any other questions? OK, so telomerase was also another important discovery that was awarded a Nobel Prize. And this gives you details. So the telomerase protects the genetic information on every cell division, though you will lose a little bit of genetic information. So it limits the number of divisions a cell can make in a lifetime.

All right, so we're going to move on now. And I've spent quite a bit of time on this, but I want to guarantee you that now, as we move forward to transcription, there's a few simplifications that we can make in the story of transcription. So moving on. All right. So what have we done so far? We've seen replication. Now we're moving to the process of transcription.

So when you transcribe something, you're basically making a copy of something, but in a slightly different format. So, for example, if you're transcribing from handwritten to a typed version, you go from something that's in the script to something that's typed. It has the same content, but it's in a different format. So this is what the process is called when you convert DNA into RNA. And very specifically, this is part of the process to make what's known as the messenger RNA.

The first phase of transcription in eukaryotic cells gets us to a pre-messenger RNA. So there's

a little bit more needs to be done to it before it can leave the nucleus to encode protein translation. But in bacteria, you're basically just going straight from the DNA to the messenger RNA. At the beginning of the next class, we will also talk about going from the pre-messenger RNA to the messenger RNA.

And let's take a look at the cell up here, where what we're focusing in on here is the process whereby we're copying that DNA. I have no idea why this is really being monstrously behaved like that. I'm done, done with these gizmos. The process whereby the double-stranded DNA opens up a little bit and we make that pre-messenger RNA copy, all right? So I want you to think back to the processes that we learned about for translation. And now we're going to move forward to take a look at transcription.

And frankly, it's a lot simpler. So let's just look at the players in transcription. And you've got a copy of this in your notes, so I don't need to necessarily put it all on the board. So in DNA, remember, we had a A, G, C, T. We have a deoxyribose, and it's mostly used as hereditary genetic information. But in RNA, we're making a new copy of the DNA, where we use slightly different building blocks-- A, G, C, U. U instead of T. Plus, there are some modified bases that occur in some of the types of RNA, and the sugar is a ribose.

So the first main thing about the RNA copy relative to the DNA copy is that ribose/deoxyribose difference. What's quite remarkable is that when you have two deoxyribose in your DNA, it's nice and stable. We need it to be nice and stable. It's our genome. We can't let our genome be falling apart as we're sort of walking down the street.

In contrast, when RNA is used, it's much more transient. We make a messenger RNA copy of part of DNA to move forward to make proteins, but we don't need that to stick around forever. And when you have the ribose with the two hydroxyls, a two and three, it's a much more fragile material. It is a transient message, and it gets degraded quite quickly. So that difference in the sugar really dictates the stability there.

RNA is found in a lot of polymers, biopolymers. We'll talk we'll focus mostly on the messenger RNA today. It's less than 1% of the DNA. And then on Friday, we'll be talking about the transfer RNA and the ribosomal RNA. So we're really going to focus in right now on the messenger RNA.

And the one thing about RNA structures I'll elaborate later is they have very different structures to canonical DNA, which adopts the double-stranded, anti-parallel structure. RNA

structures are much more like folded protein structures, where there may be sections of base pairing, but there'll also be lots of loops and different characteristics. So even the ribose structure makes a difference in the stability of the double stranded structure and encourages a lot more of these unusual structures, which is really why people have a lot of faith in the theories about the RNA world.

OK, so let's look at DNA polymerase, RNA polymerase. So here's all the good news that we'll be able to describe to you. So when you copy DNA, you copy all of it. When you copy RNA, you only copy-- when you make the copy of messenger RNA, you only copy about 1.5% of the genome. So you do not copy the entire thing. So the process is much more restricted to sections of DNA that need to be copied. And we'll talk about the features of the DNA that tell you about that later.

Here's the important details. So in eukaryotes, transcription happens in the nucleus. And the key enzyme involved is RNA polymerase, RNA pol. And it has very different features to DNA polymerase, but there are two big things that are different. It includes its own helicase. So you remember, with replication, we needed a DNA polymerase and a helicase. RNA polymerase is much smarter than that. It actually includes both functions within its structure.

So it's an RNA polymerase that grows the new nucleotide five prime to three prime. But it also has a built-in helicase, so that's an advantage. It still grows the messenger RNA five prime to three prime, but it uses the different nucleotide triphosphate building blocks. Or one of them is different. So UTP, ATP. So remember, the U replaces the T in RNA, so that's one key difference.

It includes a helicase activity. And the other really neat thing, because it's such a complication in replication, is it doesn't require a primer. That is why even when we were replicating the DNA, we were using RNA polymerase to make those little pieces of primers, because it didn't need a priming sequence. So there's really fundamental differences about the RNA polymerase.

And then the other thing is that only one of the two strands of DNA is transcribed. And in a moment, or maybe the beginning of next class, we'll judge how we can understand which sequence is transcribed. And then, obviously, the messenger RNA is a complementary sequence to the sequence of DNA that is being copied. Finally, only part of the DNA is transcribed, under unlike the process of replication.

All right, so you can see already that there are a lot of simplifications in transcription that we did not have the advantage of in replication. So the helicase activity and the primary issue are two key features that make life a lot simpler. I just wanted to show you this small detail about RNA polymerase.

There are a lot of natural products out there that are known to be inhibitors of vital processes, and one that caught my eye is the small molecules that are found in mushrooms, the really toxic mushrooms, when you see some of these. In fact, never eat a mushroom that you don't know and you know where it came from, because there's problems with them. Because a lot of these mushrooms include potent natural products.

And in fact, there's a compound known as amanitin, alpha-amanitin, and it's found in certain mushrooms known as either the Death Cap or Destroying Angel mushrooms. So you could tell from their names that they are a real problem. And what the amanitin does is it actually interferes directly with RNA polymerase by acting as an allosteric inhibitor of RNA polymerase and locking it into a closed state so it can't keep on transcribing.

So I thought this was very interesting, incredibly. Tiny, tiny, tiny, tiny doses will arrest transcription and cause dire consequences. So I think what's very interesting is that it's an allosteric inhibitor. It's very potent. What it does is it seals the polymerase in a locked, closed state that it can't move forward for transcription.

Now, finally, a couple of points. When we decide that a portion of gene is going to be transcribed, there are a lot of mechanisms in place to identify the portion of that gene. And one of the key things that is known is that there are what are called promoter sites, which are actually upstream of the portion of gene that's to be transcribed, where you recruit a bunch of proteins that actually park down on the double-stranded DNA and then, at the end of the day, recruit the RNA polymerase.

So all that extra genome, some of it is not transcribed into messenger RNA for making proteins, but it's part of an area of the gene that gets recognized by all of the proteins that collaborate to bring in the RNA polymerase in order for your RNA to be transcribed. So what I'm showing you here is a double-stranded DNA with one of the very common promoters that's just upstream of the part of the DNA that gets transcribed. And it's called the TATA box, because it's T-A-T-A sequence. It's got a complement that looks like it.

And it's shown in pink here. And then the proteins that bind to the DNA at the TATA box actually drape over that segment of double-stranded DNA and then serve as recruitment entities to bring in all the machinery that's needed for transcription of the gene beyond it. So some of the identity of all that extra double-stranded DNA is actually guide places to guide where the machinery for transcription has to park in order for messenger RNA to be formed.

So immediately, you can see we're only going to transcribe part of this genetic material beyond here. But we need a whole bunch of genetic material that's actually just serving as sort of the runway for the plane landing in the right position, all right? So I'm going to put up a puzzle that you can think about. And then we'll start with these at the beginning of the next class, because I don't want to rush them.

When you decide to transcribe a gene-- let's say you've got a promoter site here-- the thing that I want you to think about is, which strand would you transcribe? And what's the logic behind this? And then we'll just do a recap on this at the beginning of the next class, because I just want you to think about it. Because a lot of the information you need is directly here. Hi there. We're just wrapping up. OK? So that's it for today.