

**BARBARA**

So we're going to get started. This is a complicated lecture to choreograph, but I'm going to do my very, very best because I think there's some pretty amazing stuff that we have to explain that is carried out in nature. And one of those things is how do we replicate the entire genome of organisms in one fell swoop almost perfectly-- sometimes there are little errors-- and fast.

**IMPERIALI:**

So we're going to refer to these numbers in a moment. But first of all, I just want to get you back in the picture of where we were at the end of the last section, the biochemistry section. And in the next few classes we're going to address issues related to the central dogma and the use of nucleic acids for information storage and information transfer.

And before I do that, I just want to highlight a couple of things, just some terms that you should recall and recap from the biochemistry section about nucleic acids. So nucleic acids form complementary double strands through base pairing, and that base pairing involves hydrogen bonding between the nucleobases, one purine to one pyrimidine, and the AT base pair is worth two hydrogen bonds. The GC base pair is worth three hydrogen bonds.

And those are actually pretty useful facts to remember because they tell us about the stability of double-stranded DNA, where it's easier to tease it apart, and other characteristics. So it's kind of a useful thing to keep in mind. So the base pairing between what-- the purine pyrimidine pair, which sets the exact register, down the double-stranded DNA. The backbone is the exact distance apart because you always combine a small pyrimidine with a much larger purine. So that's something you should feel comfortable about.

The components are organized in an anti-parallel orientation, so one end runs 5 prime to 3 prime. The other end runs 3 prime to 5 prime. And I showed you last time an image that allows us to state quite clearly that the anti-parallel orientation is significantly more stable than the parallel orientation because you can't make all those great hydrogen bonds well in the parallel organization apart from everything else that becomes complicated.

Another important thing that you're going to have to remember, especially in this lecture, is that we add new nucleotides to the 3 prime end of a nucleic acid. And so I'm just going to put this little guy in the corner here. This is the ribose sugar. That would be where the nucleobase is attached. And it's a deoxy sugar, so there's no substituent at that carbon. This is the one position where the base is. the 2 position, where it's deoxy, the 3 position. And these are all

prime numbers. Remember, the sugars have prime numbers. The bases themselves have numbers without primes to distinguish what we're talking about. So 3, 4, 5, and they're all prime.

And when we grow the single strand of DNA, we always add new basis to the 3 prime end. That's an important numbering system. I don't like that piece of chalk. So we always grow from 5 prime to 3 prime. And that will be important to you as we go through the discussion today.

And then finally, a feature of double-stranded DNA, quite unlike proteins, which sometimes you'll thermally or with pH melt out and will often have a problem reforming their reliable structure because instead of the structure refolding, they aggregate. Double-stranded DNA doesn't aggregate because it's got all the negative charges that would be quite repulsive. So it would be very difficult for DNA strands to aggregate as such because there's too much of a concentration of negative charges. That would be repulsion amongst the same type charges.

So DNA can be peeled apart with heat, and it will reanneal faithfully to its partner a complementary strand. Once you get to six, seven, eight base pairs, that is a nice interaction between the strands, and it forms faithfully. If we have a mistake in the strands, the stability of the double strand will be a bit less, and we can measure that through physical measurements that we carry out in the lab. So we can measure the denaturation or the thermal melt temperature.

So we call this process, after you separate double-stranded DNA, we call it reannealing, or it's also designated as hybridization to form a hybrid, which is the composite of the two structures. Those terms are used fairly interchangeably.

All right. So the goals in the next four classes are to show you how the structures of the nucleic acids really are purposed for the sorts of processes that they undergo. So the replication of DNA, the conversion of a strand of DNA into complementary messenger RNA once the process of protein synthesis needs to be initiated, that second step.

The first step is called replication. The second step, when you transcribe DNA into RNA, is called transcription. And then finally, when you translate messenger RNA into proteins. So going to a completely new language, we call it a translation. And that's the basis of these lectures. There is a little bit of stuff in these lectures about more complicated issues in the mammalian cell, where we have to process the messenger RNA a bit before we can have it leave the nucleus. And I'll explain at the time when it comes. So this is the lineup for the four

lectures.

All right. Well, I already talked to you how when you add-- I mentioned it over here-- when you add cytidine triphosphate, GTP, ATP, TTP, these are the activated nucleobases, so they're the deoxy nucleoside triphosphates. They are nucleotides because they include phosphate, but when we describe them, we'll call them nucleoside triphosphates if we're going to mention the phosphate part. I know that probably doesn't make any sense at all. Is everyone OK with that?

If you're just generically describing something with a phosphate in it, you call it a nucleotide. If you want to be a little bit more specific, you say nucleoside, and then you say how many phosphates. All right? And so for the building blocks, it's a nucleoside triphosphate. And believe me, I get that wrong all the time, and my students correct me. So I don't have a problem if you're not perfect with that nomenclature.

So these are the building blocks for DNA, for polymerization. And I just showed you how you grow from the 5 prime end-- that number's not shown there, but it's shown over here-- and you add to the 3 prime end. And the convention, remember, when we describe a strand of DNA, because we make it 5 prime to 3 prime, we write it 5 prime to 3 prime just so that we're all consistent and we know what we're talking about.

Now I want to talk about two experiments that enlist the use of isotopes because these were quite useful early on to describe some of the characteristics of replication, so some of the details of replication, and also the fact that DNA was the genetic material.

So isotopes are elements that share the same number of protons and electrons, but they differ in the number of neutrons. So the common isotopes of the elements that are in all the covalent structures of the body are hydrogen, carbon, and what I'm putting as the number next to it is the common isotope, so that number designates the sum of the number of neutrons and protons. So carbon-12, phosphorus-31. I should have put nitrogen here. Let's see. Wait a minute. I'm going to put these in order because it makes more sense. Phosphorus-31, nitrogen-14, PS, phosphorus-31, and sulfur. What's the sulfur isotope? 32.

So these are the common isotopes that comprise the majority of those elements within the body. But there are lots of experiments done that have different isotopes of these atoms that we can use as traces or markers, kind of the same thing. But you'll see when I mention the word tracer, to me, it brings up the radioactive isotopes because we use very little of them. So when we talk about these elements, there are different isotopes that we use commonly.

There's the hydrogen that has an extra neutron. That's also called deuterium.

And then there's the hydrogen that is radioactive. It's metastable. It will decay and emit a radioactive particle. So this one, we would call it the heavy isotope, and this one we would call the radioactive isotope. So we could trace certain hydrogens in biomolecules by either making those hydrogens the heavier ones, or making them the radioactive ones.

Carbon also has useful isotopes. Carbon-13 is the heavy isotope and carbon-14 is the radioactive isotope. So they're used quite commonly. And then things get a little bit different for nitrogen. We use a heavy isotope of nitrogen, which is N-15, just one extra neutron. But while there is a radioactive isotope of nitrogen, it's too short of a half-life to work within the lab. It's not useful to us, so we never talk about it. But I will describe to you an experiment with the heavy N-15 when we talk about the mechanism of replication.

And then for phosphorus and sulfur, the most important ones are P-32, which is a radioactive phosphorus, and S-35, which is a radioactive sulfur. There is another radioactive phosphorus that's P-33. Has a slightly longer half life. It's kind of handy to work with if you have certain experiments. The half-lives vary a lot, but the half-life of tritium C-14 are long. That's why we do carbon dating. The half-lives of P-32 are short, less daytime frame, and sulfur 35 is a little bit longer. But those are nuances that you don't need to worry about.

Now how are these isotopes useful as traces and markers to tell us about biology and details of biology? So what I'm going to just show you here is a particular experiment that's carried out with what are known as baculovirus. So viruses infect eukaryotic cells. Baculoviruses infect bacteria. So there's a commonality there.

And when baculoviruses infect bacteria, just as the parallel with eukaryotic cells, they deposit material into the bacteria so they can replicate, so they can hijack the bacteria to make new baculoviruses. So an early experiment was done to ask what is the genetic material that transfers the information from the baculovirus to the bacterial cell in order to make more baculovirus.

So they were able to demonstrate, in this Hershey and Chase experiment, that you could label protein and DNA with particular isotopes and find out what part of the baculovirus was important for transferring information for the production of new baculovirus. So they wanted to label the capsid protein, so the proteinaceous material that's on the outside of this lunar

lander. This is a baculovirus. It's sort of amazing how it looks.

So they wanted to label the protein. But they also wanted to label the contents of the baculovirus, the DNA, with tracer radioisotopes. So what would be the best isotopes to use if you wanted to differentiate between protein and nucleic acid?

And what you want to think about is, what does protein have that nucleic acids don't have, and what do nucleic acids have that proteins don't have? Which are the elements that are important for differentiating? Yeah. Over there.

**AUDIENCE:** Phosphorus.

**BARBARA** Yes, phosphorus. Do you want to give a little bit more of a--

**IMPERIALI:**

**AUDIENCE:** [INAUDIBLE]

**BARBARA** OK. So the answer is we'd use phosphorus to label nucleic acids because that phosphodiester backbone is rich in phosphorus. There are a few phosphates in proteins, but they're very transient. They're part of signaling. But every nucleotide has a phosphorus in every building block. And then in proteins, they have sulfur in them where nucleic acids do not. And the sulfur is in two amino acids cystine and methionine. So you can use those two tracers for the building blocks.

And the way the experiment works is if you've labeled the protein with sulfur, you infect a cell, the virus replicates. And what you're going to do is then centrifuge the bacterial cells to see what's in the cell, and then you can, alternatively, label the DNA-- and in this case, I've shown it as green, but that would be the other isotope-- let it infect the cell and deposit its contents in the cell, centrifuge the cells out. And you want to know where the radioactivity is, and the radioactivity is strictly associated with phosphorus-32 because the genetic material that coded for the production of new baculovirus is the thing that stays associated with the bacterial cell.

In contrast, there's no radioactivity associated when you label the cells with S-35. All right? Nice experiment. Easy way to do it. Rather nice way to do it.

The other experiment I want to describe to you very briefly because it relies on central irrigation technology that's very powerful. And it's a method that utilizes N-15 very, very nicely. So let me tell you about that experiment. There are things in the laboratory that we use every

day, ultra centrifuges and regular centrifuges, that spin at a very fast speed where we can really differentiate things by molecular mass, which relates to sedimentation coefficients. How fast does the particle spin to the bottom of the tube when it's under high centrifugal force? The heavier it is, the faster it will spin.

So the question that came up very early on-- it looks like a nonsense question now because it looks so obvious that we replicate DNA the way we do. But originally, it wasn't absolutely certain whether DNA was replicated through a semi-conservative mechanism, where the strands came apart and you made two copies, a copy of each strand, to make two identical daughter new double-stranded DNA.

Alternatively, could you have a mechanism that was quite conservative, you kept your DNA, and somehow you made a copy of it. A little harder for me to understand. So your two new strands, one would look like the original one, but one would be very different. Or dispersive, where you're just copying bits of the DNA and somehow reassembling this gigantic jigsaw puzzle.

The centrifugation experiment allowed people to absolutely and clearly state that the replication is via a semi-conservative process. So let's walk through it. And what was done is that the nucleotides within the DNA were all labeled with heavy nitrogen. So for every nucleic acid-- let's say, for example, when there's an adenine in the backbone, there are five nitrogens, so that would be five atomic mass units heavier than nitrogen-14 in that nucleobase.

So N-15 nucleobase, five mass units heavier than N-14 nucleobase. Make sense to everybody? All right. So it's heavier by some amount. It's not a massive amount, but it's enough to differentiate in a centrifugation experiment. So the bacteria were first grown up in heavy nitrogen, so all of the DNA sediments at a certain rate and place. And it's all got N-15 in it, so it's as heavy as it can sediment.

And then they let the bacteria replicate in the presence of N-14, the lighter isotope of nitrogen. And what one would anticipate to happen is whenever you replicate, you peel apart the two heavy strands, and each one pairs with a light strand. So you now have two copies, two new identical copies of DNA where half of one strand has N-15, one strand has N-14. So it will sediment less quickly because it is of a different density, a lighter density than the all N-15. So you would get an intermediate weight band in the centrifuge tube when you're sedimenting.

If you peel those two strands apart again, you've got a heavy and a light. If you're growing in N-14, the light will combine with another light, and there are two of those. And then the heavy will combine with a light. So you'll get new sedimentation where you still have a band that's the combination of the light and heavy. But then you start having some material that is all of the light. So this would be this cycle. So two that are still light plus heavy, and then two that are exclusively light. You keep on replicating, and you'll keep on diluting the heavy. Does that experiment make sense?

It's kind of a cool experiment. You can hardly believe it's feasible, but the centrifugation ability to differentiate with these isotopes is really valuable. So I just want to put in a plug for the use of isotopes. Obviously, they're used in nuclear physics. We use a lot of the radioactive isotopes for different reasons. But in biology, they are indispensable for some of these experiments. And even to this day, you can do things with isotopes you can't by other experiments, either the radioactive or the heavy.

There's a great deal of work done nowadays in proteomics using mass spectrometry and heavy isotopes, where you can really track where things go and treat cancer cells differently from healthy cells, but actually track what's happening by putting heavy isotopes in one of the growing dishes of cells.

So I like both of those experiments a lot. I would put them in the category of oldies but goodies.

All right. Now I've got some details over here that we've got some explaining to do as we move forward. And I just want to, first of all, start with a couple of details that I want to highlight on this board.

First of all, prokaryotes such as bacteria, E. Coli, and eukaryotes such as human cells. Have some differences in their DNA. So the size of the bacterial genome is about five million base pairs. The size of the human genome is not quite 1,000 times bigger, but a good deal bigger. The DNA in bacteria is circular, whereas the DNA in eukaryotic cells is linear.

And it's actually in the form-- you see all these little images where you see the pairs of chromosomes sort of stuck together at the center, and that's linear DNA from one end to the other. In bacteria, the DNA is circular, so it doesn't have an end. So they're different, all right?

So replicating circular DNA is a little bit different from replicating linear DNA. Now in both

cases, the DNA has to be packaged up so it will fit in the cell. Otherwise, it's just too much of a disordered thing. In the case of bacteria, the circular DNA is wrapped up and super coiled with what are known as polyamines, compounds that are very positively charged to neutralize all that negative charge in the DNA. And in man and other eukaryotes, the chromosomes are wrapped around very, very positively charged proteins known as histones. They're similar principles, but they're different entities.

So when we start talking about replication of a prokaryotic cell, here's the typical circular DNA. So you can see the double-stranded DNA. And I've put a little symbol here that I call the ORI. So that's origin of initiation. So that is the place where you start copying your DNA. And we're going to talk about how to spot these places in genomes in a moment.

And so the bacterial genome is copied bidirectionally, so you can go in both directions copying in the appropriate direction-- we'll talk about that in a moment-- to make the entire circular DNA and a perfect copy of the DNA with bidirectional copying. And that's not bad for small genomes because you've only got one origin of replication. You've got to get all the way around the circular DNA. It's quite a bit smaller than the human one.

But what do you do with gigantic genomes such as the one that's about 1,000 times bigger than the bacterial one? How do you manage to still catalyze the replication of the entire genome quickly enough to make the copy of DNA in some reasonable amount of time? And that is taking also into consideration that the speed of bacterial replication is 1,000 base pairs per second. Pretty impressive-- 1,000 of those bonds made every second-- whereas in eukaryotes, it sits somewhere down from 50 base pairs per second, depending on conditions. These are not very explicit numbers. There's a little bit of variability.

So how would you do it? If you had this massive chunk of gene, and you've got to copy it just really pretty quickly in order to replicate an entire cell contents in eight hours, what would you do? How could you expedite things? Yeah.

Yes. So the answer here is start in a lot of different places so there's a lot of collaborative work going on all along that large genome. And so when the replication of the long linear chromosomes that are in eukaryotic cells, you end up with just a lot of origins of replication. So you're basically starting all over the place so you can get the job done quickly enough for it to make sense. Does that make sense to everybody? So start in a lot of places. It's a pretty easy thing.



All right. Now I mentioned very briefly that before you can copy DNA, you have to unpack it. Now we think a lot about packaging DNA, less about unpacking it. So I've got these pictures sort of in a reverse direction. So it's only this form of DNA stretched out that can be copied, whereas the DNA in your cells is bundled up very tightly into chromosomes.

How are they put together? Those chromosomes are made up of chromatin, which is a bunch of balls which are histone proteins with DNA wrapped around them. That's compact chromatin. Sorry, guys. Let me go back, not to go forwards. So this is chromatin in a compact form. This is chromatin now unraveled. They look more like beads on a string, DNA wrapped around each histone protein to form those nucleosomes structures.

And then each nucleosome looks like this. It's got protein in the middle and DNA wrapped around it. So in order to go forward and replicate, we have to unpackage DNA. There are lots of signals to unpackage the DNA that we will talk about later. The nucleosomes look like this. So you can trace the DNA and the proteins bundled in the middle.

And if we have time at the end, I will show you the video of the packaging of DNA and the video of the replication of DNA. They're about a minute and a half each, so it's not a whole bunch.

Later on, we will talk about what happens when there is a determination that a cell needs to replicate its genome. A bunch of things have to happen as signals to do the unwrapping. And one of those things is actually to alter the charged state of the histone proteins so you neutralize the positive charges so the DNA can unravel for it. Makes sense. It's still just a plain old electrostatic interaction.

So the synthesis of DNA is what's known as template-driven polymerization. OK. So template-driven polymerization. You'll get a sense of exactly what that means. We know the two strands are complementary, and we're going to use one strand as the code to make a new complementary strand. So when we're making DNA, we have one strand there. That's what's known as the template strand. I want to make a complementary copy of that strand.

So what the DNA polymerase is going to do is systematically add nucleotides to the 3 prime OH of the ribose, and then you keep growing. But the base that gets put in is the one that's complementary to the base on the template strand. So if it's thiamine on the template strand, it's adenine put in on the new strand, and so on.

And just to get these numbers again, the new strand is grown from 5 prime to 3 prime. The old strand is read 3 prime to 5 prime. So you're reading a strand and you're reading it 3 to 5, and you're making 5 to 3. That's how you end up with complementary anti-parallel strands. OK with that, everyone?

All right. And when you form the bond, you go in with nucleoside triphosphates, you form a new phosphodiester with just one of those phosphoruses, and you kick out phosphate phosphate, which we would call pyrophosphate or diphosphate.

OK. So here you could just sort of see, just check that you know what you're doing. I'm going to just say this is really a pretty simple thing. We're going to put in a T, then we're going to put in an A then we're going to put in a C because our template is telling us to put in the appropriate purine or pyrimidine that forms a nice base pair with that.

So there's a question here. But I've already given you the answer to it. So it should be very straightforward for you to look at a template strand and decide what would be on the complementary strand, and also decide on the directionality of the complementary strand.

Now origins of replication. What do they look like? It's the old where do I start problem. Where do I start? What's the best place to start? Someone? Not you. Not you. Anyone else want to answer here? I've got this whole genome. I've got to get going. I've got to start making a copy of it. Where's the best place to break into it to start making copies? Yeah. There.

**AUDIENCE:** [INAUDIBLE]

**BARBARA IMPERIALI:** Correct. It's simple. It's the difference between two and three. If you go to the area where there's lots of G's and C's, every GC has three hydrogen bonds. If you go to a different area where there happen to be a lot of A's and T's in a row, each pair is only worth two hydrogen bonds. You're going to pick the stretch that has the most A's and T's because it's the place where the base pairing is weaker.

It's actually very simple concept. So it would be the AT-rich region is the place where origins of replication are more predominant. OK. So what we're going to do now-- and this is going to be tricky, but I'm going to make this work-- is we're going to talk about replicating an entire chunk of gene. And what I've done here is I've put the pieces that we're going to discuss over here. That's my menu. And we're going to work through what we need to copy a large chunk of DNA.

All right? Everybody with me?

So I redrew this this morning. I promised myself I'm going to be super tidy on this blackboard, which is a far stretch for me. So replication, we know where replication starts, AT-rich region. We could pick those out in the genome. We know that we've got to unwind the DNA to make a copy of it, so we must have had to unwrap the DNA beforehand. These are all known things. So let's now take a look at double-stranded DNA and try and figure out how we use all the components.

Actually, I'm going to bring this board down because I don't see it as well up there. How do we make use of all these components that are part of the menu that we need to make the new strand of DNA? So we have double-stranded DNA and there are base pairs across. That's typical double-stranded DNA. That's what we're starting from, unwrapped. I haven't made it helical because my helices will get messy.

And so we need to start. So the first thing that will happen is that the important enzymes like the helicase will scan to find an origin of replication, so somewhere for the process to start because that's where they're going to break in.

So David, one of your TA's, works in Steve Bell's lab. And they study the mammalian origin of replication complex, which is this much more complicated situation than what I'm going to describe to you today. But they do fabulous single molecule work to show how those pieces are assembled.

But I'll tell you the truth. As dorky as I am, the thing that I find the most cool about the origin of replication complex is that it spells ORC. And if any of you are *Lord of the Rings* fans, who can't be excited by a complex called the ORC complex?

So there's going to be the screening of the genome for ORI. And I want you to remember what is up here is the rest of the chromosome. It's not just a fuzzy end. It's something that's got a lot of stuff there that's still base paired.

So the first thing that happens is that helicase needs to start unzipping the double-stranded DNA, so you get to a new intermediate. We're going to draw like this.

And the helicase is going to intercept to basically start separating those two strands of DNA in order to do the replication. But don't forget that these are still base paired, and these now are

single strands. So I'm going to need something from that menu. I'm going to need it quickly.

What do you think's the next thing we're really going to need in order to be able to move forward in our job? What do we need from over here? We're already using the double-stranded DNA. It's not going to be too difficult to figure out the NTP's. We're already using the unzip arrays, the helicase. What are we going to need? Yeah.

Primase. Ah, primase shortly. But we've really got to do something to stabilize our unzipped stuff.

Anyone else?

So those two single strands are right near each other, and they're complementary. What's going to stop them going straight back together again and not allowing replication? So there are proteins that basically sit on the single-stranded parts of DNA as are known a single strand binding proteins that stabilize this transiently made complex long enough for it to be copied.

So so far, we've used the helicase. We've used the single strand binding proteins. And I'm going to draw them there, still remembering that I have an entire chromosome up here. Now the heavy lifting has to start fairly soon, so we need the enzyme that's going to start making a copy of one of those single strands.

That's going to be a DNA polymerase, which is right here. But DNA polymerase is kind of a finicky enzyme because it doesn't like to just start cold on the genome, on the single strand, and start making a copy. DNA polymerase needs a primer of some kind. So I'm going to just put that as a note up here. DNA polymerase needs what's called a primer.

And what is a primer? A primer is a sequence-- I think I've got something up here-- a primer is a sequence that is complementary to a little bit of DNA so that when DNA polymerase comes along, it's actually grabbing a double strand and then filling up the rest of the single strand. So this would be a typical depiction of a primer. Here's a strand we want to copy. But the blue one is a primer strand giving you something double-stranded to hold onto.

And then in this situation, DNA polymerase would be happy to fill in the rest of the sequence. So for the purpose of this discussion to start, we're just going to provide a primer that is complementary to the DNA just to give us a go. We have an alternative in the cell, where it's to actually use a primase and RNA building blocks to build a primer. But let's keep things simple for a minute and get on with a lot of the major work. And then we'll come back to that in a

second.

So this part is a primer. And that primer-- if this is the 3 prime end, that's 5 prime 3 prime. So the primer is complementary to the DNA in that direction. It's anti-parallel. And then DNA polymerase can go along and fill in the strand. And I'm going to draw it dashed, and it's going to keep on growing.

And what I want you to notice is DNA polymerase grows from 5 prime to 3 prime. That is a cardinal rule, the directionality that we grow the new DNA in. So basically, looking at this picture, it's going in this direction because the 3 prime OH is free, and you're building onto it. So you use a primer first, and then you use DNA polymerase plus deoxy nucleoside triphosphates. So all those AT CG building blocks.

So we've been using these. We've used the single-stranded binding proteins. We've used a primer, but there are an alternative. I'll get to that in just a second. But we now have a complex where I'm going to briefly divert you to tell you about, in the cell, what else you could use as a primer. And then we're going to have to deal with copying this other strand.

But this first strand that is made is called the leading strand. It's the one that's easy. I peel apart the DNA. I have a primer there, and I can just build. Goes really smoothly. I'm building in the right direction. If you look over on the other side of the fence, you've got a problem because I can't build on this side from here, right? Why not?

**AUDIENCE:** [INAUDIBLE]

**BARBARA IMPERIALI:** Yes. I'd be doing it the wrong direction. I would be the total mess, frankly. There would be a crisis in the cell. So we're going to have to deal with that. So let me just get you out of the primer, just deal with this primer issue. In the cell, it's not like we can micro inject all these little primers to get our DNA synthesized.

So RNA polymerase, believe it or not, doesn't need a primer. So you can build little bits of primer with RNA. You don't build very much, just build short segments. And then once you've got sort of enough there, DNA polymerase can start polymerizing.

So you use RNA polymerase and nucleoside triphosphates, not the deoxy ones. So later on, we've got this piece here that might be made deoxy RNA, which is why, coming over here to my menu-- do I have it? Yes. We need an RNA primase, and we need RNA's to chop them out

so that DNA polymerase can come back because now there is double-stranded material and fill in the gap. And then we need one more enzyme to stitch together the gaps. And that's the ligase. The logic of this stuff is great. I think it's just really amazing because all these moving parts have evolved to make all these functions fall into place.

Now we need to talk about this pesky other strand. So we have a problem here because we've got the wrong directionality. So what happens in nature is as soon as there's enough of a stretch, a primer is put in place, growing the DNA in the right direction. That's from where the helicase is unzipping. So we put in another primer. And then DNA polymerase-- that's the white one-- can build its complementary strand, building in the appropriate direction to be consistent.

So you basically fill in a chunk and then you have to wait awhile till more is unzipped, put in another primer, and then grow the DNA. So in that other piece you're making little segments of DNA intervened by RNA primers. Later on, those small primers are going to be chewed up by RNA's. DNA polymerase is going to fill them in, and ligase is going to join them. Does that make sense? It's a lot to grasp.

And this other side is called what's known as the lagging strand. And there are pieces of short DNA that are transiently made that have their own name. You can remember it or not. I just feel I'm not being complete if I don't mention their name. They're called Okazaki fragments after the guy who discovered them. And they're longer in bacteria than in humans. So Okazaki fragments are the transient short stretches of DNA that are made in the lagging strand, and then they get zipped together.

So that's sort of the story for replicating the DNA. Let me make sure I've described everything to you. Leading strand, lagging strand. Oh, can't believe I almost forgot this. All right. Now we got a problem. So we've come all this way. We've used all the pieces. We haven't used one of the pieces.

So we're in a small bit of trouble because what's going to happen when you're, for example, trying to peel apart your DNA? What trouble are you going to run into? Who wants to come up here and get involved in my demo? And you've got to be aware that you're going to be on screen. You and you up there. You've got your hand up first. OK.

I need someone who's going to hold onto this as if their life depends on it. You can't let the yellow ones slip out. So you are the rest of the chromosome, so make sure that you really

hold-- and you don't let things come apart.

Now you are helicase, OK? OK. So come on. Start helicasing and really pull as hard as you can. He's going to hold it. No, I don't want you to unwind it because we're in the middle of a chromosome. So you've literally got to do what helicase does. Yeah, keep on pulling. Come on, you can do it. Can you go much further? I mean, your arms aren't any longer. But it's almost impossible, right? We get to a stage where we need help.

The help we have coming is from topoisomerase. Topoisomerase does something-- so walk just one step this way. Yeah. Topoisomerase some comes along and it cuts the DNA. It holds the pieces in its hands, if you will, lets the super coiling relax, and then it joins them back again. So topoisomerase-- thank you very much, guys. That's it. Chromosome, helicase, thank you. We really have some tightly wound DNA here.

So topoisomerase is the get out of jail free card because it allows you to deal with all of this tension you got in place that you cannot unravel. So you need someone who is literally going to cut, hold, let the thing relax, rejoin it. And there are some topoisomerases called DNA gyrases.

But the really cool thing about topoisomerase is that it's a wonderful drug target in both mammalian and bacterial cells because they're quite different in the two types of organisms. So in bacteria the antibiotic ciprofloxacin is a topoisomerase inhibitor that actually stops your bacteria dividing because if topo doesn't work, you can't go on.

And in human biology and cancer biology, there are also topoisomerases-- there's one called [INAUDIBLE] -- that does the same thing in the eukaryotic topoisomerase to stop cancer cells dividing rapidly so that those cannot go on and make larger tumors.

So I want to show you something. And I am actually certain that I have time. Come on. Now I'm going to show you.

[VIDEO PLAYBACK]

- In this animation, we'll see the remarkable way our DNA is tightly packed up so that six feet of this long molecule fits into the microscopic nucleus of every cell.

The process starts when DNA is wrapped around special protein molecules called histones. The combined loop of DNA and protein is called a nucleosome.

Next, the nucleosomes are packaged into a thread. The end result is a fiber known as chromatin. This fiber is then looped and coiled yet again, leading finally to the familiar shapes known as chromosomes, which can be seen in the nucleus of dividing cells.

**BARBARA  
IMPERIALI:**

Now I'm going to take you forward to DNA, what we just talked about here.

- Using computer animation based on molecular research, we are now able to see how DNA is actually copied in living cells. You are looking at an assembly line of amazing miniature biochemical machines that are pulling apart the DNA double helix and cranking out a copy of each strand.

The DNA to be copied enters the production line from bottom left. The whirling blue molecular machine is called helicase. It spins the DNA as fast as a jet engine as it unwinds the double helix into two strands. One strand is copied continuously and can be seen spooling off to the right.

Things are not so simple for the other strand because it must be copied backwards. It is drawn out repeatedly in loops and copied one section at a time. The end result is two new DNA molecules.

[END PLAYBACK]

**BARBARA  
IMPERIALI:**

All right. OK. That's what you saw today, replication, a lot of the mechanics, a lot of the moving parts. There's this tricky stuff with primers. You'll get used to that. But this is really the quintessential set of pieces to replicate DNA.

It is an amazing process. Look at the speed. 1,000 base pairs a second. Can you even believe it? An entire circular chromosome copied in 20 minutes. So these are really things to think about because they are so impressive that it's a delight to actually be able to teach them to you.

OK. I'm done for today.