Biochemistry for one lecture 35
Today we're going to talk about replication. We're going to talk about the leading strands, the lagging strands, and the difference between replication of bacterial and eukaryotic systems. We're going to talk about the enzymes and key proteins involved, and about regulation, and then we're going to finish with a discussion of topoisomerases, topoisomerase I and II. So let's get started.

Replication involves the synthesis of a complete genome in preparation for cell division. Now as you can imagine, during the replication of the genome initiation of the process is highly regulated. This must occur at the right place and at the right time. The initiation of replication occurs at specific sequences in the DNA and it occurs only one time per cell cycle, and this is to make sure that we don't have additional copies of DNA being produced in addition to the normal genome.

Now there are differences and similarities between the replication in bacterial and eukaryotic organisms. The bacterial genome is one circular molecule, consisting of 4.8 million base pairs. Cell division occurs based on cell size, and is thought to be linked to an availability of nutrients. For the eukaryotic genome, humans have 46 linear chromosomes, with 6 billion base pairs in the genome. That's a lot to copy. Cell division takes hours, and it is highly regulated. Cell division occurs in response to intracellular signals and extracellular signals, and it's affected by cell type and developmental state. Now as we can see from the figures below, the bacterial genome has one origin of replication. The DNA is unwound in one place and replication starts there. It proceeds bi-directionally, going along the right-hand side of the circle, and the left-hand side of the circle, and meeting in the middle to terminate. Eukaryotes have several origins of replication per chromosome in order to get such a large amount of DNA copied before cell division.

Initiation occurs at the origin of replication. This is abbreviated ori. In the bacterium, *E. coli*, this is known as *OriC*. This is a specialized region of the DNA that consists of about 245 base pairs. It contains two main regions. The first is a tandem array of three AT-rich consensus sequences that are high in adenine and thymine, and this consensus sequence is shown in green. Now for every AT base pair there, there are two rather than three hydrogen bonds that are holding these two nucleotides together, and so regions of high AT content are easier to separate. This region also contains DnaA protein binding sites.

DnaA is an initiation protein that helps to open the strands at the origin of replication. It has an intrinsic ATPase, and when DnaA is bound to ATP, binds the DnaA sites with a higher affinity than if it were bound to ADP And so what happens is the DNA is going to bind to the *oriC* and it's going to unwind the DNA beginning at the AT- rich region, but the thing is, is once the strands are separated, it's important to keep them separated, and not let them bind again to one another, and so the DNA becomes coated with SSB protein. SSB protein is single strand binding protein.

DnaA and also helps to form the pre-priming complex and a way to remember what the pre-priming complex is in E. coli is to remember ABC. DnaA unwinds the DNA, and recruits DnaB, and DnaC. Now DnaB is a helicase that is a hexameric ring, and it becomes bound to the lagging strand. And the way that this happens, is this, DnaA helps to recruit DnaC. Now DnaC is a carrier protein that carries DnaB to the site of initiation, and DnaC is guided to this site by association with DnaA. And so DnaC binds in this region and helps to load DnaB onto the lagging strand. Now DnaB is not only a helicase that unwinds the DNA, it's also a platform for DnaG, and DnaG is a primase that's going to synthesize the RNA primers for the lagging strand. And together, DnaB and DnaG form the primosome. The primosome has directionality. DnaB faces forward in the direction of the replication fork. We see it pointing to the left on one side and to the right on the other. Now that the primosome has been formed at the site of initiation, it's time to recruit the main polymerase. This is the polymerase that's going to synthesize DNA.

And so replication in prokaryotes, which is very similar many ways to that in eukaryotes, goes like this. A DNA helicase separates the DNA into single strands and the DNA template is read three prime to five prime. RNA primers are polymerized first, five prime to three prime, by DnaG primase, and then the main replicative polymerase in *E. coli* is going to polymerize the DNA, five prime to three prime. And this polymerase is the same both in the leading strand and in the lagging strand. The other thing is, a hetero-duplex between RNA and DNA is not as stable as a DNA-DNA duplex. We all know that RNA can be hydrolyzed more easily than DNA, and so it's necessary to remove the RNA primer. DNA Pol I is the enzyme in *E. coli* that removes this primer, and as it does, it adds DNA in a five prime to three prime manner, and so the thing is, with DNA polymerase

one, in some ways, is like a snow plow, removing the nucleotides ahead of it one by one, and as it does, it adds DNA. Once DNA polymerase has removed the RNA primer, and has added the DNA to replace that primer, there's a gap there, and DNA ligase is the enzyme that seals the gap.

During replication, the DNA is synthesized five prime to three prime according to the template, which is read three prime to five prime. Now there are two replication forks for every origin of replication, one to the left and one to the right. In this diagram, the direction of the replication fork is shown by a (dark) blue arrow. The top strand is the leading strand, and is being synthesize five prime to three prime, in the same direction as the replication fork. However, the lagging strand is being synthesized five prime to three prime in the opposite direction, and so to accommodate this backwards direction, this DNA is actually synthesized in short fragments known as Okasaki fragments, and these are shown in (light) blue.

And so DNA polymerization is a complex job and requires a complex machine. This is the machine that we see here. It's called DNA polymerase. The whole polymerase consists of the clamp loader, two beta-2 siding clamps and core polymerases. The replisome consists of the beta-2 sliding clamp, and the core polymerase. The core polymerase itself in *E. coli* consists of an alpha polymerase unit, which actually is the catalytic region, epsilon subunit, which is a three prime to five prime exonuclease, and a theta subunit. It is thought that this subunit offers some sort of a structural stability to the core polymerase. The beta-2 sliding clamp is a donut-shaped protein that consists of a dimer of beta proteins. The opening in the center of the doughnut is just wide enough to accept double-stranded DNA. The sliding clamp associates with the core polymerase to hold it on the DNA going down the strand. The polymerase continues to hold onto the substrate, which is the template DNA, without falling off, and this characteristic is called processivity. Now the thing is, is that in order to synthesize the discontinuous strand of the lagging strand, the clamp loader must release the polymerase and then attach it as the polymerase follows the replication fork, as the polymerase moves further down the DNA. Now, if you'll notice, there are two replisomes, this is because there are two core polymerases. One that is associated with the beta-2 sliding clamp on the leading strand, and one that's associated with the beta-2 sliding clamp on the lagging strand. In this