Cell Biology: Cellular Organization, Division and Processes

Prof. Shikha Laloraya

Department of Biochemistry

Indian Institute of Science – Bangalore

Lecture 11

DNA Replication-Part II

Hello everyone. Welcome back to the second part of the lecture on DNA replication. In this lecture we will discuss the components of the replication machinery and the mechanism of DNA replication.

(refer time: 00:38)

So, just to recap the key factors which are required for DNA replication in prokaryotes are mentioned here. First of all we need an initiator, a protein termed DnaA that binds to the origin DNA and melts the AT-rich region opening up the helix. A helicase the protein DnaB which unbinds DNA at the replication form to generate single-stranded DNA template strands. A primase the DnaG protein that synthesizes the RNA primer, which is required for DNA synthesis by DNA Polymerase.

Of course, most importantly we also need the DNA Pol III holoenzyme: this has the clamp loader, this is a part which loads another protein known as the sliding clamp onto a primer template duplex. And there is the tau protein, there are three numbers of this which connects the clamp loader to the DNA Pol III core and also contacts and stimulates the DNA helicase. And the DNA Pol III core enzyme which is of course required for the DNA synthesis reaction and the sliding clamp which is a protein complex that clamps the DNA Pol III core onto the DNA and it enhances the processivity of the polymerase.

There are also single strand DNA binding protein-they bind to single stranded DNA near the replication fork and they prevent the annealing of these template DNA strands to each other or annealing upon themselves. And some other enzymes include DNA Pol 1, which is required for removal of the primers in the end and filling the gap and DNA ligase, which is required for ligation of Okazaki fragments that are short DNA fragments that are formed during lagging strand DNA synthesis.

One also needs Topoisomerases; in *E. coli* these are Topo 1 and Gyrase and the function of these enzymes is to prevent entangling of DNA during replication and they remove super coils on the DNA. And in fact there are some other factors also which help in making the replication more efficient or regulated in some ways, which are not mentioned here. And these are all you might have noticed many of them have a name starting with Dna and they came out of a screen using conditional

mutants again to look for factors that were defective in DNA synthesis. So, the main ones are shown here.

(refer time: 04:12)

So, the requirements for replication of DNA in *E. coli* are mentioned on this slide and of course the *E. coli* genome as you know is a circular genome and it has got a single origin of DNA replication, which is referred to as oriC. Now just the size of the DNA of *E. coli* is 4.6 times 10 to the sixth nucleotide pairs (4.6×10^6). And the replication forks progress at approximately 500 to 1000 nucleotides per second. The initiation step of DNA replication in bacteria particularly *E. coli* is regulated when the conditions are optimal then replication is initiated. And replication initiates when the initiator protein DnaA binds in multiple copies to specific sites in the origin and it creates a stretch of single stranded DNA nearby. The DNA helicase binds to this complex and then it gets loaded onto the adjacent DNA single strand.

The DNA primase binds the helicase and it forms a complex referred to as a primosome, which synthesizes an RNA primer in order to initiate the DNA synthesis.

(refer time: 06:08)

Shown here is a structural basis for initiation of replication by the DnaA initiator protein. DnaA forms a long helical assembly of subunits. And it is thought to perform two important tasks, initially its double-stranded DNA binding domain binds to the DNA double helix and it wraps it around the outside. So, that is shown here- the initiator protein binding to the double helix.

And it forms a sort of a super helical structure next to the sequence which is rich in AT base pairs and that region which is rich in AT base pairs is of course more weakly bonded. So, then the DnaA binds to this AT-rich region and it helps to melt the double helix by capturing and extending one of the DNA strands. As seen in this structure here the DNA strand which has been captured and extended is shown in orange.

(Video Start Time: 07:28)

As you all know DNA is a double helix, for replication to occur each strand is used as a template and for copying the genetic information on the template strand the helix must be unwound to expose the bases and allow the polymerase enzyme to access and incorporate the complementary nucleotides. So, unwinding of the DNA is harder than one might expect, the interaction between the bases is quite strong.

Even though it is via hydrogen bonds, which are relatively weak - they are 20-fold weaker than the covalent bonds, there are many of them hence it is it takes appreciable energy to separate these strands. And this function of course is performed by DNA helicase in cells. And these are enzymes that pull apart and unwind the two strands in the DNA double helix.

(Video End Time: 08:25)

(refer time: 08:26)

So, we will go through the steps in the replication of DNA sequentially and I hope we will be able to appreciate the process of DNA replication. So, after the initiator DnaA has captured the single stranded DNA at the origin, DNA helicase or DnaB protein binds to this complex with the help of another protein the DNA helicase loader DnaC, which was not mentioned in my list which I showed you earlier.

The DNA helicase loader helps in opening up the DNA helicase protein ring and placing it around the single stranded DNA at the origin. And the DNA helicase now gets loaded onto the DNA single strand adjacent to the DnaA. The helicase then recruits DNA primase that binds to the helicase forming the primosome, which synthesizes the RNA primer which is required to initiate DNA synthesis by DNA Polymerase. The primer causes the release of the helicase loader from the complex.

(refer time: 09:55)

Shown here is a structure of DnaB helicase, which is shown in blue colour- the subunits are in blue and it is in a complex with the helicase loader which is shown in magenta, and parts of the primase protein are also here in green. So, what I want you to note is couple of things, one is the hole in the middle through which the DNA can pass. And also I would like you to note the gap in the ring of the helicase over here, which could be the site of entry of the single stranded DNA.

(refer time: 10:41)

Shown here is the structure of a bacterial helicase, the helicase itself is shown in blue colour and it is a hexamer it has got 6 subunits. And it is encircling DNA which is shown in orange. Helicases use ATP, ATP is depicted here in red, to push DNA through a ring of protein subunits. So, here there are six identical subunits arranged in a ring and this ring is not flat, it is somewhat helical like a lock washer shape.

The six subunits they follow the DNA helix inside the hole in the center that they are surrounding. And the bacterial helicases are more or less similar to RecA whereas eukaryotic helicases are triple A plus ATPases. So, both the types both of them they have got a ring of subunits that surrounds the single standard DNA. And as the helicase surrounds this single stranded portion of the DNA which is shown here in orange, it actually makes its way along the single strand and separates the double helix as it goes along and it uses ATP to power this motion.

(refer time: 10:41)

So, this is just a different representation of the same structure where you can actually see the path of the single standard DNA more clearly. So, each of the subunits here is shown in a different colour

each is a different polypeptide chain. And the single stranded DNA is in the middle in the dark green or teal colour. I hope you can appreciate this. So, you can note the spiral path of the DNA in both the images.

The second image has been rotated a little uh sort of like this, to show the path of the DNA from the side view. So, here again you can see the spiral or helical path of the single stranded DNA, which is trapped inside the ring of the helicase. So, in this structure it actually revealed that the DnaB helicase adopts a closed spiral staircase quaternary structure around this ssDNA single stranded DNA. And its C-terminal domain coordinates two nucleotides of the ssDNA and it might use a hand over hand mechanism where there is sequential hydrolysis of NTPs and this causes a sequential 5 prime to 3 prime movement of the subunits along the helical axis of the spiral staircase. And this results in unwinding of two nucleotides per subunit. Below, shown are the top and sideways and the bottom views of the same molecule. Because this is a different representation do not try to match the colours between them.

(refer time: 14:53)

So far, we have learned that the unwinding of the parental duplex DNA is carried out by the replicative DNA helicase that couples NTP hydrolysis to 5 prime to 3 prime translocation. Because the leading strand synthesis starts first and it is continuous and efficient, there's hardly any single stranded DNA template accumulating on that side where the leading strand is. So, the annealing of the two strands of the duplex DNA to each other is less likely.

However, the single stranded template DNA formed by unwinding the duplex DNA by the helicase could also self-anneal. And in order to prevent this there are proteins known as single strand DNA binding proteins abbreviated as SSBs that bind to that single-stranded DNA and they prevent it from annealing to itself. The majority of these proteins, they have got a distinct structure or topology.

And this is known as the OB domain- the oligosaccharide or oligonucleotide binding domain. This domain forms a group that surrounds a DNA and structurally it is shared between SSBs from different organisms and this domain is also important for the oligomerization of the protein. So, each of these single strand DNA binding proteins also binds to other DNA bound single strand binding proteins and thus it forms cooperatively long rows of SSBs, which straighten out this unwound template DNA and this further facilitates DNA Polymerization.

The SSBs bound to DNA, they can be easily displaced though by the DNA Polymerase. So, their presence does not hinder the process of replication per se.

(refer time: 17:21)

Now the primase binds to the helicase and it lays down an RNA primer and forms this important structure the primer template junction, which is then recognized by another protein complex, the clamp loader. The clamp loader recognizes the helicase and the primers particularly the primer

template junction. And this clamp loader is a complex which is part of the DNA Pol III holoenzyme and it binds to a sliding clamp and this is a ring-like structure.

And by binding to it, it creates a gap in the ring-like structure so, that the DNA can enter and be encircled by this clamp. The clamp loader encircles a new primer template duplex and it carries the clamp along with it till it reaches a three prime end of the primer. And then hydrolyzes ATP and releases the clamp and while it does that it allows the clamp to close. So, actually the cells use these clamps to help make processes more efficient.

For example, the cells they tether their DNA based molecular machine such as this polymerase to the ring shape clamp. The clamp can slide freely along the DNA and it keeps the machinery of DNA synthesis on the job by preventing it from diffusing away if it falls off and it facilitates its rebinding by hanging on to it. And this arrangement leads to remarkable increases in efficiency by tethering the DNA Polymerases to the sliding clamp the bacteria increase the speed of replication sometimes hundred-fold and also increase processivity. The processivity means the number of bases that are copied before the polymerase falls off and must start again. So, processivity definitely is enhanced by the sliding clamp.

(refer time: 20:21)

Two conformations of the clamp, open versus closed, would be shown here. The first one is with the open clamp. The clamp loader is shown in blue, it grabs on to a clamp, which is represented in green and it opens it up and allows the double stranded nucleic acid to bind inside. The nucleic acid here is part of the template which is represented in pink, paired with the prime,r which is represented as the orange strand. ATP is used to power the very last step that is to release the final clamp-DNA complex.

(refer time: 21:13)

So, closing of the ring can be seen in the second conformation. So, this is the open conformation and here you can see the closed conformation. So, this happens when the clamp loader along with the clamp reaches the end of the 3 prime primer. Now the clamp loader can be released and the clamp encircles the primer template duplex completely.

(refer time: 21:43)

DNA Polymerase III is a replicative DNA helicase in *E. coli* and as already mentioned it synthesizes the DNA by adding nucleotides to the 3 prime hydroxyl of the primer and it is got a single active site for catalysing the DNA synthesis. And it has very specific properties that it will incorporate only the correctly paired bases as specified by the complementary sequence of the template strand.

And it is got a unique structure which resembles a partially closed hand where the DNA fits in a large cleft. It is highly processive and for example the *E. coli* Pol III can add up to 1000 nucleotides per

second. And the holoenzyme, the DNA Pol III holoenzyme has got three core Polymerases and also a clamp loader and three tau subunits which are attached to flexible linkers and the clamp loader.

(refer time: 22:57)

Shown is a cartoon of the DNA Pol III holo enzyme. And the DNA pol III holoenzyme consists of three copies of the Pol III core enzyme and one copy of the sliding clamp loader. The sliding clamp loader has got three copies of the tau proteins shown here in red each of which interacts with the Pol III core enzyme. The tau region is connected to the clamp loader by these flexible linkers.

And this linker allows flexibility to allow the Polymerases to move in an independent manner while being associated. And it is required for one to replicate the leading and the other two to replicate the lagging strands. So, it allows a lot of movement while the polymerases are doing their job. The clamp loader itself has got 5 subunits and they are arranged in a helical arrangement, and it can hold the clamp in this open lock washer shape.

(refer time: 24:25)

Shown here is the structure of the budding yeast replicative pol delta complexed with a primer template and the PCNA clamp that was deduced by electron microscopy. So, not bacterial now, this is the budding yeast one that I am showing just for the sake of example. This 3.2 angstrom resolution cryo EM structure was deduced by Zheng et al and it shows the budding yeast Pol delta in a complex with primed DNA strand which is shown in green and the PCNA clamp which is this multicolor multi subunit ring in the bottom.

And I hope you can see the nucleic acid strand in the middle. The light green is shorter and is supposed to represent the primer and the darker green represents the template. So, one of the interesting features that was revealed was that Pol delta was binding only to one of the subunits of the PCNA trimer. And this singular interaction with only one of the subunits is extensive and it holds the DNA quite well such that the two nanometer wide DNA threads through the centre of this three nanometer interior channel of the clamp without directly contacting the protein. There is a water mediated interface between the clamp and the DNA that passes through it. This interface allows the PCNA clamp to water skate, as per the authors, along the duplex with minimum amount of drag. So, it can easily slide along. There is also cartoon shown on the left; do not try to match the colours but it shows the polymerase, the clamp, and the DNA.

(refer time: 26:51)

Now after closing and releasing the clamp the released clamp loader remains in the vicinity of the polymerase. In order to assemble a new clamp on the lagging strand at the start of a new Okazaki fragment. A sliding clamp is assembled on each primer and leading strand synthesis is initiated by one of the core polymerases of the holo enzyme, the clamp also binds the Polymerase.

The clamp can slide freely along the double helix that it encircles; the sliding clamp binds and holds the polymerase. So, that even if the polymerase gets released from the primer template junction it helps it reassociate without floating away and thus it enhances the processivity of the enzyme as it incorporates the nucleotides. After the helicase has progressed about a thousand bases or so another primer, a second primer, is synthesized on the lagging strand and again a sliding clamp is loaded onto that.

The resulting primer template junction is recognized by the second Pol III core enzyme of the holocomplex and lagging strand DNA synthesis is initiated. The first Okazaki fragment is then extended by the second Pol III core enzyme and simultaneously more lagging strand single stranded DNA template is generated on which a new RNA primer is also synthesized. So, now there's another primer-so, another sliding clamp is assembled by the clamp loader. And the third Pol III core enzyme now also participates in lagging strand synthesis to make the second Okazaki fragment making the process more efficient. So, now the fork has assembled with the replisome and this process repeats and continues.

(refer time: 29:16)

Shown here is the replisome at the replication fork and also an explanation of the trombone model for coordinating replication at the lagging strand. The steps are that the helicase on the lagging strand template it moves in the 5 prime to 3 prime direction unwinding the duplex DNA and creating these 2 single stranded template strands. The DNA polymerase 3 holoenzyme interacts with the helicase via these tau subunits and stimulates the rate of strand separation. (This particular interaction is not depicted in this diagram though). Periodically the primase enzyme associates with the helicase and it synthesizes a primer on the lagging strand template. The primer is depicted in yellow colour and it is important to remember that in order to notice it in the upcoming animation.

(refer time: 30:37)

The helicase extrudes a single stranded DNA loop and the loop is initially small but it increases in size and shown here is the loop in the largest possible configuration. The single stranded DNA binding proteins bind to the single stranded regions of the loop to prevent its self annealing. The lagging stranded DNA is folded to bring the lagging strand polymerase in a complex with the leading strand polymerase molecule and also the three prime end of a completed Okazaki fragment close to the start site for the next Okazaki fragment. Since the lagging strand Polymerase is in a complex with the other replication proteins, the DNA Polymerase holoenzyme, it can be reused to make successive Okazaki fragments, and of course for this it would be alternating with the third Polymerase core enzyme which is shown here as the unengaged one right now.

Now the size of the single stranded loop between the helicase and the Polymerase on the lagging strand changes. It grows in size and it is then when it reaches the primer of the next Okazaki fragment it is released. Then a new one is formed and that also grows in size. So, this keeps on

repeating and that is why this model is referred to as a trombone model. The trombone as you may know is a trumpet-like instrument that has got a sliding mechanism that varies the length of the instrument to change the pitch-this extends the length of the air column and that lowers the pitch.

A clamp is put at the primer template junction by the clamp loader. The clamp can also bind and hold the polymerase core on the lagging strand in place and it enhances its processivity. The polymerase core enzyme extends the chain at the end of the primer initiating the Okazaki fragment. When the enzyme has reached the end that is the primer of the last Okazaki fragment then it can no longer extend there's no further to go. So, then it lets go and the Polymerase is now released from the sliding clamp. However, before this one lets go, another primer was synthesized and the Polymerase core III which is shown as the unengaged Polymerase right now had bound and it starts extending it. So, now the one which was making the Okazaki fragment earlier becomes unengaged from the DNA strand and lets go, while the third one is extending and forming a new Okazaki fragment and this process continues in succession till the process of replication has been completed.

(refer time: 34:23)

Now with this background knowledge let us again see the events coordinated at the replication fork by the replisome in this animation by Drew Berry and Etsuko Uno.

(Video Start Time: 34:37)

In this animation the helicase is shown as a bright blue structure unwinding the double-stranded DNA at the left side. And you can see the leading strand being synthesized continuously at the bottom; the lagging strand is shown at the top. So, the greyish green primase arrives and binds to the helicase and it forms a primer (in yellow). The clamp loader, which is depicted in lavender, recognizes the primer template junction (in yellow) and loads the clamp, which is the green ring around the primer template junction.

The lagging strand Pol III core (in purple) binds and initiates the synthesis of an Okazaki fragment. Note the loop predicted by the trombone model as it increases in size with the progress of the lagging strand replication. Now here is another loop being formed and it is increasing in size. When the polymerase reaches the end of the earlier Okazaki fragment it lets go and it is available to re-initiate at another primer template junction.

The clamp which is there prevents the polymerase from falling and floating off while it is synthesizing the DNA. So, these clamps they hold the enzyme in place as was explained earlier, you can see this right here. Unlike the explanation that I gave you, here only two polymerase molecules are shown, not three. That is there is only one for the lagging strand, not two as explained earlier, and this is probably because when this animation was made it was not known that there are three polymerase molecules in the DNA Pol III holoenzyme.

I hope this depiction helps you understand the dynamic and complex events at the replication fork mediated by the replisome.

(Video End Time: 36:52)

(refer time: 36:53)

There are some intriguing ways by which the components of the replication machinery cooperate to enhance efficiency. For example, the primase activity is stimulated by binding to the DNA helicase at the replication fork. It binds to the helicase to initiate a RNA primer and it is then released. Initiation of the next primer requires rebinding of the primase through the helicase. And the dependency ensures that the primer is made only at the replication fork and not elsewhere at other single stranded regions in the genome.

Another interesting point is that the DNA Polymerase III holoenzyme contacts helicase via the tau subunits of the clamp loader and stimulates its unwinding activity. Also, the sliding clamp enhances processivity of the replicative DNA Pol III enzyme. So, different components of the replisome, they sort of cooperate to enhance the efficiency of the process.

(refer time: 38:09)

Now coming to the end, and finishing off replication.

There is a final step in completing the lagging strand replication. The RNA primers, which were formed on the lagging strand they have to be removed to complete the DNA replication and this is done by DNA Pol 1 in *E. coli*. Removal with these primers leaves a gap that is also filled by the DNA polymerase and the remaining nick is sealed by a DNA ligase thus completing the lagging strand DNA synthesis.

Termination of replication occurs at the converging forks; the replication forks that were initiated at the origin oriC in a replication bubble they move in opposite directions away from each other. The fork progression stops when they collide head-on with another replication fork moving in the opposite direction, or, when they reach the end of a chromosome in case of linear chromosomes or eukaryotes.

So, this is how the replication would end. Now there is a problem in eukaryotes that occurs because they have linear chromosomes. The replication of the end requires a special mechanism in linear chromosomes; this is referred to as the end replication problem. During the lagging strand replication, a backstitching mechanism is employed and this results in the problem that when the fork reaches an end of the linear chromosome there is no place to produce the RNA primer needed to initiate the last Okazaki fragment at the end of the linear DNA molecule. So, to address this, eukaryotes have repetitive sequences at the ends of their chromosomes forming telomeres that recruit an enzyme known as telomerase. Telomerase extends the repetitive sequences on the parental DNA strand and the replication of the lagging strand at the end can be completed by using

these extensions as a template for the synthesis of the complementary strand by the DNA polymerase.

So, I hope this lecture has clarified aspects of the mechanism of the mysterious process of DNA replication. Thank you.