Organic Chemistry In Biology And Drug Development

Prof. Amit Basak
Department of Chemistry
Indian Institute of Technology, Kharagpur

Lecture - 32
Central Dogma: DNA Replication, Transcription and Translation (Contd.)

Welcome back to this course on Organic Chemistry in Biology and Drug Development. In the last session we have discussed the process of translation. Before that we studied the process of transcription. Again just to have a quick glance of the information flow in biological system - at first the DNA is self-copied by semi conservative replication i.e. from double stranded DNA you get a copy DNA. Then the DNA has to be transcribed to a messenger RNA.

The other RNAs are involved in the subsequent steps i.e. translation. So, transcription is the process of copying of the DNA into the RNA. Then from the RNA protein will be synthesized according to the sequence of codon inscribed in the RNA. The messenger RNA is particular type of RNA that gives the primary structure of the protein that will be synthesized.

So, replication transcription and translation are the three fundamental processes of living system. If, any one of these can be stopped, then the cell division that is mitosis that cannot take place. So, the cell will die automatically. Now, just a quick glance of these three fundamental processes, remember in replication always try to find out what are the enzymes. Because, we will always look at the organic chemistry perspective of the processes involved in replication.

Now, in replication there is lot of enzymes or proteins. Some proteins which are involved have no enzymatic activity but binding have affinity. Now, the first thing you have to break the helix so, a helicase enzyme is involved. Then you have DNA polymerase which will synthesize the polymer. Then you have the primer which are made up of RNA.

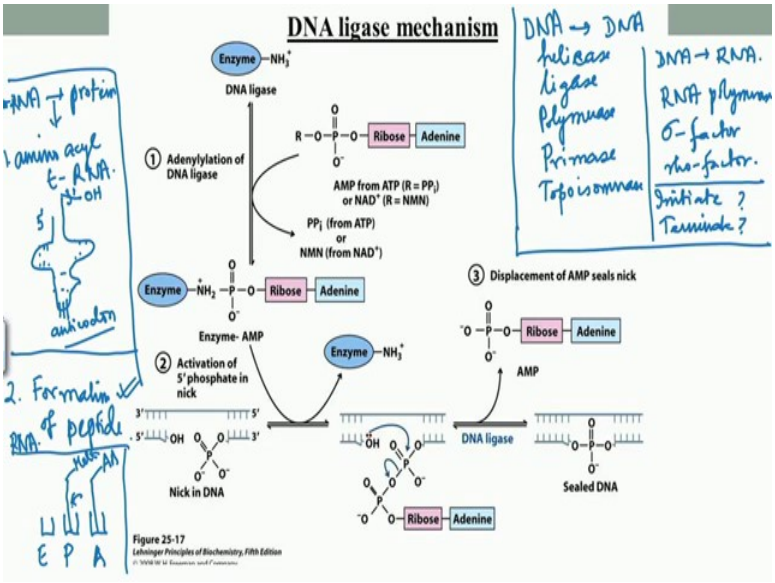
RNA polymerase makes short primer which is made up of ribonucleotides only. Then you have the generation of the Okazaki fragment. There is a problem in going from the 3 prime to 5 prime directions. Because the DNA synthesis takes place in the 5'-3'

direction. So, it is a discontinuous process and you get fragments which are called Okazaki fragments.

So, you need an endonuclease enzyme which cleaves RNA primers and then that is replaced by proper polymeric short polymeric oligonucleotides. After that there is ligation process where you have to join these Okazaki fragments via an enzyme called DNA ligase.

There is this SSB proteins i.e. Single Strand Binding proteins whose job is to keep the two strands apart because your DNA polymerase is working to make the double strand.

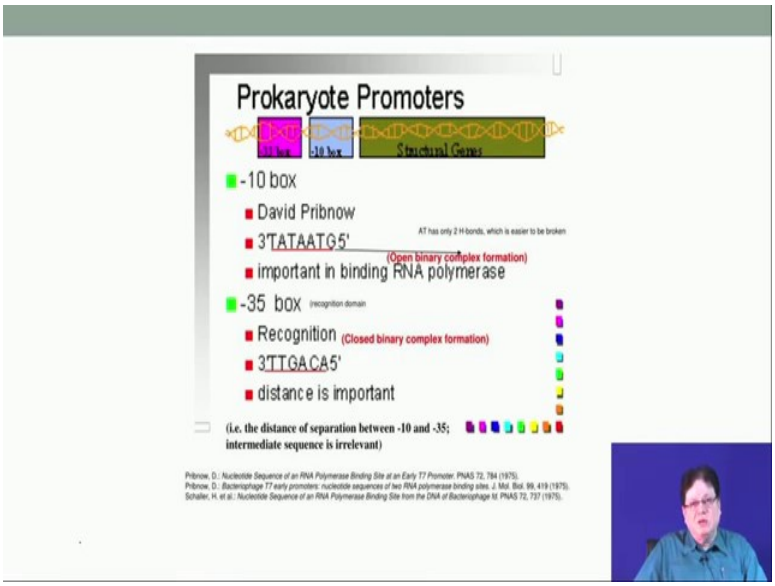
(Refer Slide Time: 04:33)



Many enzymes and also non enzymatic proteins like SSB proteins are involved. When you unwind the DNA there is a strain that develops in the front of the replication fork and that is super coiling. It is taken care of by an enzyme called DNA topoisomerase, because it is basically a topological problem. If it disturbs the geometrical properties ahead of it that is called a topological problem. So, you have a topological problem which is generated by supercoiling.

In the transcription, RNA polymerase binds at the promoter region. I told you about different promoter regions, one is the Pribnow box in prokaryotes. We are talking about prokaryotes.

(Refer Slide Time: 06:39)

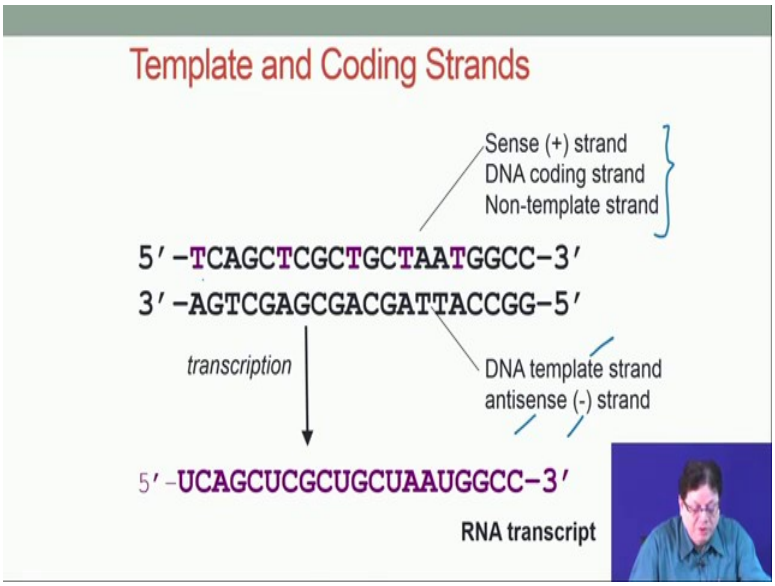


You have the minus 10 box which is called the Pribnow box. The sequence is from 3 prime to 5 prime. The sequence TATAAT was obtained from particular type of bacterial strain. If this sequence is present at the Pribnow box then there will be attachment of the RNA polymerase with the strands of the DNA. Remember there are again two strands- one is a coding strand, another is a non-coding strand. One is sense another is antisense or you can also call it a template strand and a non-template strand.

The binding to the template strand will be will be very high. If there is a sequence perfect homology is there in the bacteria. It is the Pribnow box, then you have the minus 35 box. Remember this is actually the upstream. So, if you have the first base of the RNA that is considered to be the plus one and this is transcribed. We are talking about this upstream. If you write it from the 5 prime to 3 prime it will be on the left side.

That is called the minus 35 box and there is again a consensus sequence. But again I remind you that it may not be perfectly followed by all bacteria. So, after the binding, there is this RNA polymerase. It is a pentameric protein. Then there is a sigma factor and that sigma factor is the one which is actually recognizes this sequence. Then binds to the DNA strand and then the sigma factor falls off. The RNA polymerase moves downstream and bringing one base or the other.

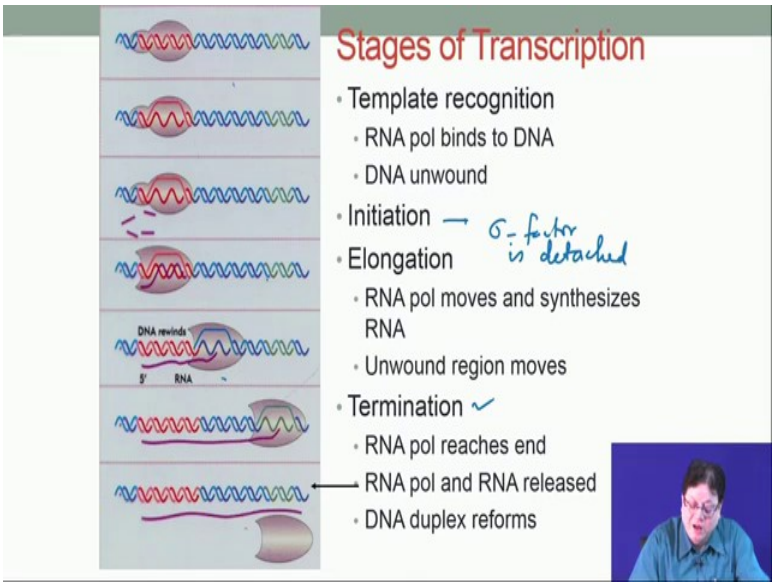
(Refer Slide Time: 09:35)



This is again the same thing and I just repeat what is your sense strand. These 3 are the names that are given to this. This is basically the DNA template or anti sense. Sometimes it is also positive strand or negative strand. That is also another terminology called i.e. non template strand.

You have to be very careful when we say something. So, all these different names for the same strand and different books will utilize different names. Here this is antisense strand and this is the sense strand. Why is it sense? Because, this is the sequence which will be transcribed into the RNA. The two differences between RNA and DNA are- there is U in RNA instead of T and the sugar is ribose instead of deoxyribose.

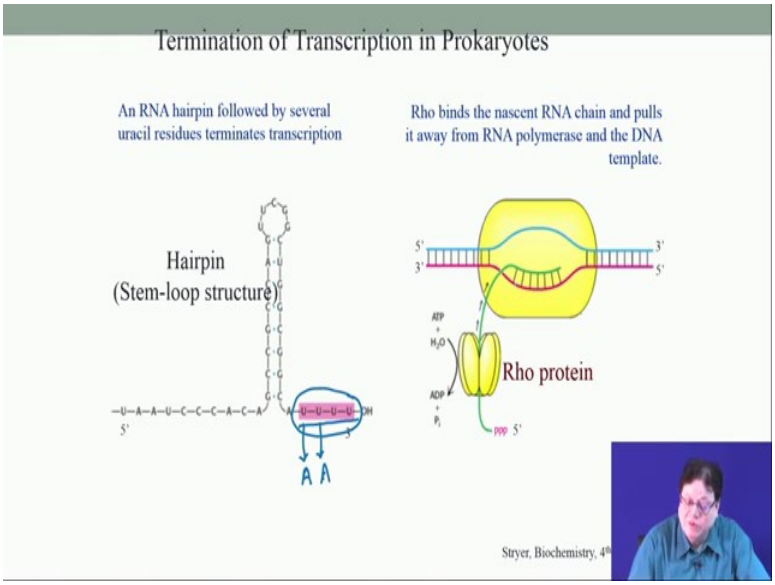
(Refer Slide Time: 10:43)



So, I think these are all done last time. This is sigma factor and it falls off. Now, this sigma factor is not there. So, the RNA polymerase can move bringing in the oligonucleotide bases. This synthesis of oligonucleotides goes from the 5 prime to 3 prime direction. Once it is done it reaches the termination step. Then everything falls off.

There are the different steps- template recognition, RNA polymerase binds to DNA, unwind, initiation, elongation and then termination. Here, it is just initiation and then basically the sigma factor is detached. It can move and synthesize the complete RNA. Termination process is basically- it reaches particular type of geometry and then it falls off at that time.

(Refer Slide Time: 12:11)



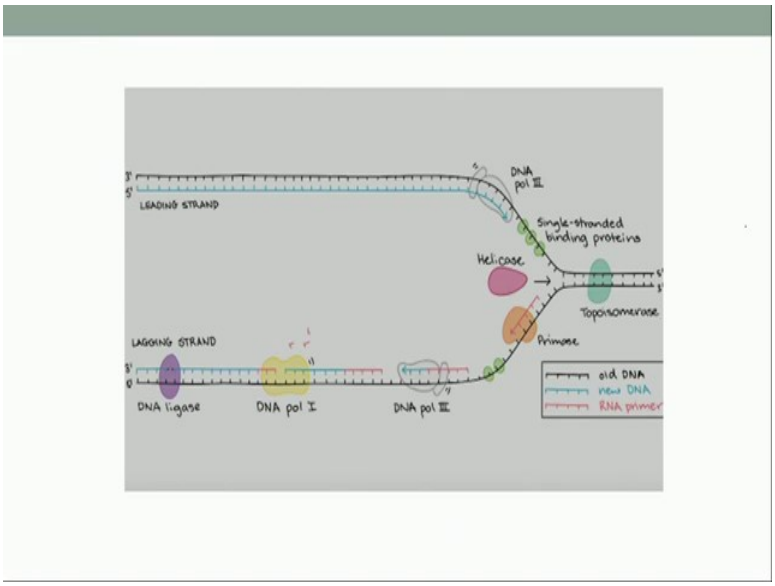
There is a GC rich region which forms a hairpin. The hairpin is basically poly U. There are so many Us. These are basically the U that U complements with A. That is the only 2 hydrogen bond forming process and that part will be very loose.

So, that can easily melt i.e the separation can be easy. So, as it melts it falls off from the DNA and then the whole RNA comes out. The other is that there is a rho dependent process. This rho dependent process utilizes ATP, utilizes the energy from the ATP. It hydrolyzes, takes the energy and then it releases. The energy is utilized in releasing the RNA from the nascent RNA and from the template DNA strand.

So, these are the different processes. In the replication, we have DNA. For DNA to DNA synthesis the required enzymes are- helicase, ligase, polymerase, topoisomerase.

The process for synthesizing DNA to RNA is called transcription. RNA polymerase is involved here and it is actually a pentameric protein. So, you should know what a sigma factor is. What is a rho factor? So, usually these are the row factor.

(Refer Slide Time: 15:17)



Translation is the process in which you now make protein from mRNA.

Here the enzymes aminoacyl tRNA synthatase incorporates the amino acid as the ester where it is hooked to the 3 prime overhang. Overhang means you have an extra, which is not connected with each other. Earlier there was definitely this type of complementarities. There are 3 bases which are called anticodon. So, what happens? This 3 prime OH is attached to an amino acid via an ester bond. It is hooked to the carboxylic functionality of the amino acid. How many tRNA are involved in peptide synthesis? The minimum number has to be 20.

Because, there are 20 amino acids that has to be brought in to the ribosome called i.e. the protein making factory. You need 20. Because, there are at least 20 tRNAs. Some of the codons can be degenerate. There are number of codons which codes for the same amino acid. 64 anticodons are possible here.

Some of the anticodons gives signal to stop the process. If it is stop codon they do not bring any anticodon. That means, there is no amino acid attached to the 3 prime end. So, the one enzyme which incorporates this amino acid to the 3 prime end that is the aminoacyl tRNA synthatase. It makes the OCO, then the R and NH_2.

When the mRNA first binds to a a region in the rRNA ribosomal RNA it is recognized. The messenger RNA comes recognizes the ribosomal sequence. Then it sits in their

ribosomes which have 2 subunits. One is a small subunit and another is a large subunit. That mRNA recognizes the sequence in the rRNA and then it spreads along the ribosome.

Then there are 3 sites in the ribosome. I told you one is called the Exit site, another is the P site, and the third one is A site. A is the amino acid site, P is the peptide site, and E is the exit site. So, what happens there? The starting amino acid is methionine. Fortunately methionine has only one codon and that is AU starting from 5 prime to 3 prime AUG.

So, the methionine containing the tRNA that will bind to the P site. Then in the A site the next amino acid as dictated by the codon next codon that aminoacyl tRNA will bind ok. So, very simply we can have these three sites, this is the exit, this is the peptide site and this is the amino acid side. So, initially tRNA is containing methionine and this is dictated by the codon sequence.

The methionine is transferred here and the dipeptide is formed. The ribosome moves and then this tRNA is now devoid of any amino acid because the methionine is transferred here. So, that goes to the exit site. And, finally, it goes away from the system. Now this peptide as it shifts one frame.

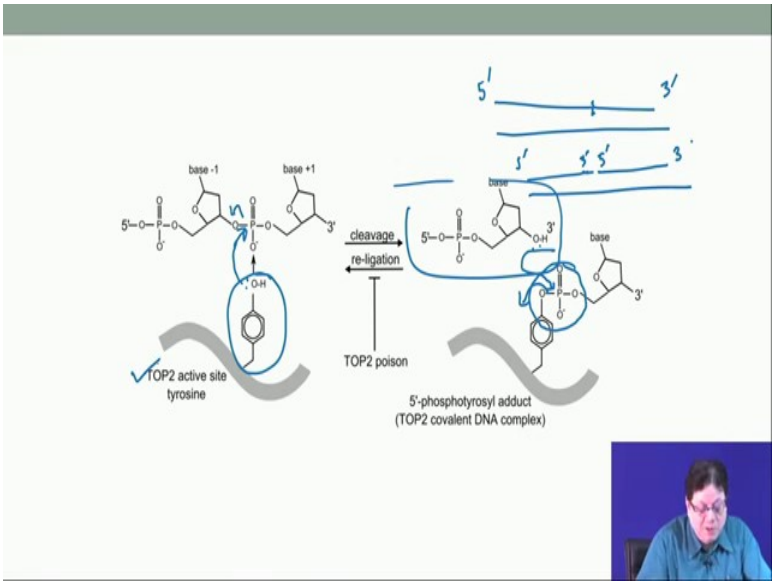
Whatever is attached here that will come in this site and this will be free and here the codon is present in the mRNA. Corresponding tRNA will bring the next amino acid. Then the tripeptide will be formed and this process follows till a stop codon comes and then everything is released. The peptide bond is released. Now, this is a reaction which is the peptide bond formation.

The formation of peptide is catalyzed by the RNA. See, there are splicing of introns. When for eukaryotes the messenger RNA is formed it is first bio synthesized. That is called the immature or pre mRNA. Pre mRNA means when there is this intron and exon both are present one after another. Then what happens? Then, there is the enzyme which is called spliceosome. Spliceosome is basically the RNA itself and it takes care of the intron, cuts the intron away and joins the exons one after another.

So, takes the intron away and joins the exons. It is catalyzed by the RNA itself. It is a very remarkable discovery because all that breaks down. It is not always true that bio

catalysts are only consisting of proteins. There are RNA molecules which can also catalyze reactions and these are called ribozymes.

(Refer Slide Time: 24:05)



When the replication fork moves a stress is generated ahead of the replication fork.

When there is stress the corresponding strain will be involved. Then to release that stress or strain you have to rotate the DNA because there is excessive super coiling ahead of the replication fork. So, you have to release that stress and that is done by the DNA topoisomerase.

(Refer Slide Time: 25:19)

Suppose you have DNA with right handed helical pattern. Now you try to unwind the system.

Now the number of helix is spread out in a very short region. It is very difficult to move because it creates a stress ahead of the replication fork.

There is a strand which is above. You cut it and then you bring it at the bottom point and then reseal it. You just hold it for the moment till the replication fork is near to this. That means, the DNA topoisomerase works by a nick. Formation of a nick in one strand helps the strand to rotate and then rotate in the opposite direction releasing the stress that is generated. The DNA synthesis is going on the polymerase.

As it comes near to that point the nicking has to be sealed because the DNA polymerase will move forward. So, there cannot be any nick. The DNA topoisomerase goes further ahead of the replication fork and then again there is the supercoiling that happens there. So, it does again nicking there and then release the stress.

There are different this topoisomerases - topoisomerases 1, 2, 3, 4. We are not going into that details. We are just showing one mechanism by which the nicking of one strand is carried out. Remember nicking means you are breaking a particular strand.

See this is the 5 prime phosphate and this is attached to other sugar base and this is the phosphodiester. This is the phosphodiester and this is the next sugar base system. Basically we are saying that the strand has to be cleaved.

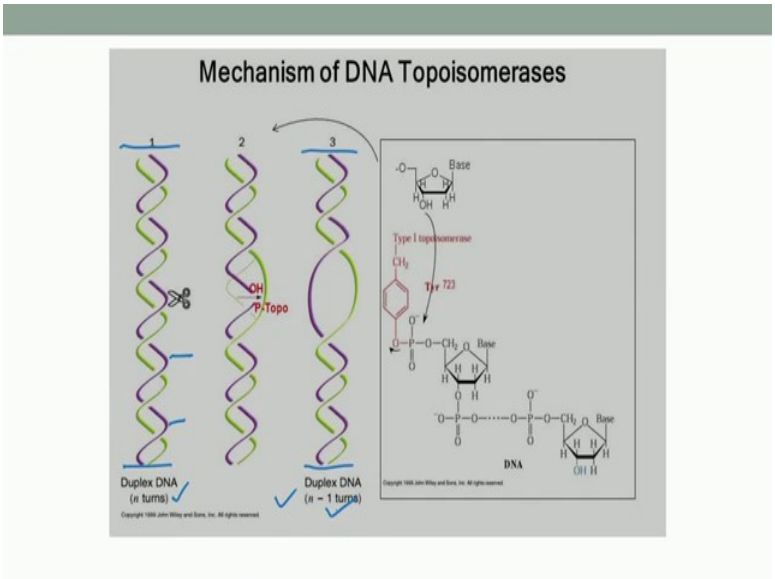
So, these are the double strand. One strand needs a nick i.e. it is not joined like this. This can turn around and release the strain. This nicking is carried out by just hydrolyzing the phosphodiester bond. If you hydrolyze the phosphodiester bond there is this no joining between the 2 ends here. This is the 5 prime end and this is the 3 prime end.

There is tyrosine in the active site of topoisomerase 2. Tyrosine is nothing but an aromatic amino acid with a para hydroxy group. This OH can be a good nucleophile and attacks the phosphorus.

This bond breaks towards the 5 prime end, this is the 5 prime end. So, this becomes OH and this is the phosphodiester. It is momentarily formed between the tyrosine and the sugar phosphate backbone.

There is a nicking now. There is no joining between this part and the other part. The turning takes place and that stress is released and the fork can proceed further. When you require the religation; that means this is cleavage and you require the religation. This O H will attacks this phosphate and the tyrosine is released.

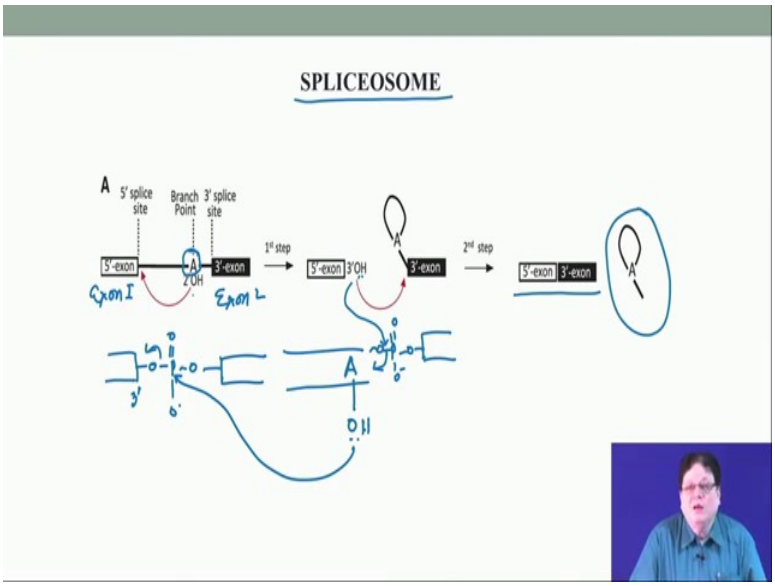
(Refer Slide Time: 31:19)



So, it is a reversible reaction. You can do this phosphodiester cleavage. So, that is one of the mechanisms of the topoisomerase 2. So, it is a tyrosine mediated phosphodiester linkage. As the replication fork moves, you have this supercoiling. So, topoisomerase is like a scissor. The scissor cuts one of the strength and then the strand rotates. It is a replication bubble.

If you have n number of base you will have (n-1) number of turn in the duplex DNA. You have one less turn in the duplex DNA. Less turn means less coiling.

(Refer Slide Time: 32:35)



Let us talk about the spliceosome. Spliceosome takes care of the intron problem. The pre mRNA or the immature mRNA goes to the mature mRNA. Here the RNA itself is the catalyst.

This is the A i.e. adenine. These are not deoxy systems. These are only the normal ribose. You will have the 2 prime OH. This 2 prime OH attacks. This phosphate linkage is attached to the last base of the exon. Suppose this is exon 1, this is your exon 2 and this is intron 1.

So, this A is residing in the intron and the 2 prime OH is attacking the phosphate linkage between the terminal base of the exon 1 connected to the base sugar of your intron. That phosphodiester linkage is cleaved. You will get a system where this 3 prime OH will be free. This is the exon. You have OPOO minus O and then this is the intron. So, there is a A in the intron and that A has 2' OH. This attacks that one and releases this one.

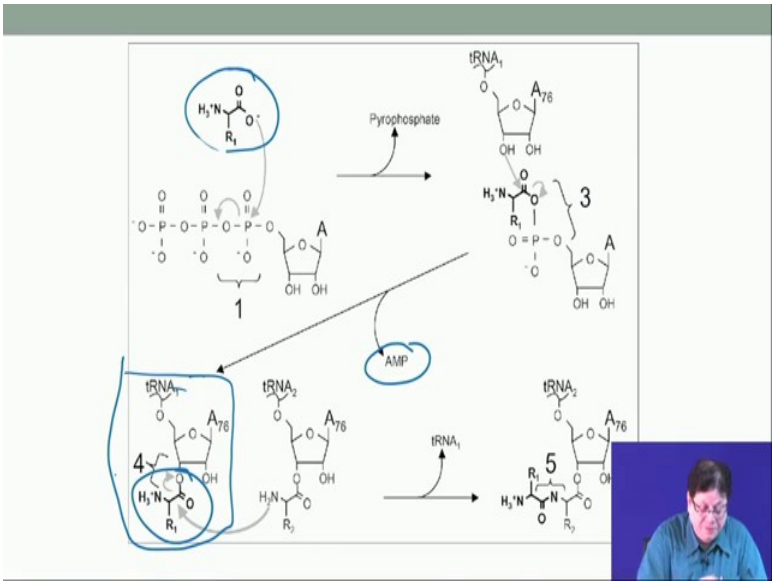
So, you get a 3' OH. The 3 prime OH is free and in the process you get this circular type intron. But it has got some attachment to the 3 prime end. This is now the 3' OH which is generated. It acts as a nucleophile and cleaves the phosphodiester linkage here.

So, this is O P O O and this is O minus. The OH which is generated here attacks this phosphate and releases the intron. So, the intron is released and the exon one and exon 2 are joined. So, that is one of the mechanisms of the spliceosome. It is basically internally

done. Why the 2 prime OH is a nucleophile? Adenine or guanine have imidazole ring. Imidazole is attached to a purine ring and these are quite basic.

The purine bases activate the 2' OH by abstracting by this hydrogen. When this hydrogen is required to be delivered it delivers it. So, that is the mechanism of spliceosome catalysis.

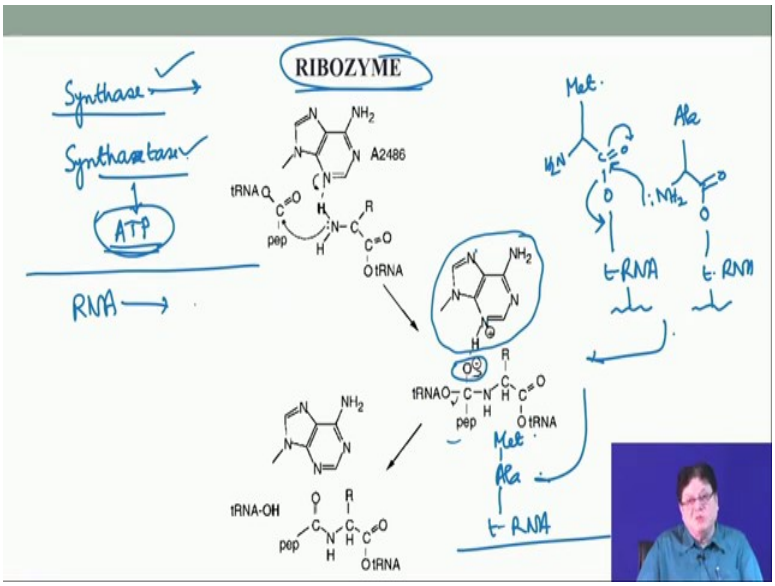
(Refer Slide Time: 36:27)



The formation of this aminoacyl tRNA is done by the amino acid and ATP molecule. ATP triphosphate. The carboxylic end is activated as the phosphate. So, first this is aminoacyl AMP, because you have a mono phosphate. Now, this 3' OH attacks this the carbonyl end of the carboxyl ester.

So, basically this is a mixed anhydride of the phosphoric acid and the amino acid. This mixed anhydride is a very much activated system. So, that 3' OH attacks this carbonyl and releases the AMP. So, AMP is released and this is transferred to the 3' OH. This is the tRNA containing the aminoacyl group.

(Refer Slide Time: 37:51)



So, this is basically the tRNA aminoacyl synthatase. Is it is called synthatase because it requires an ATP. Remember there are two types enzyme systems who does synthesis-synthase and synthatase. Synthatase utilizes ATP as a cofactor. If the synthesis does not require ATP it is called a synthase. Now, this is the mechanism of the peptide formation. At first there is a methionine.

The NH_2 is free and this will be attached to the tRNA. The tRNA contains an anticodon which is specific for the codon required for expression of methionine. The other tRNA is attached via the 3' prime group. Suppose second one is alanine and this is a NH_2. This attacks the carbonyl and that goes out.

So, the tRNA containing a methionine becomes free and this is transferred to the next one. So now, you will have a tRNA which is attached to your alanine, and then methionine. Then there is shift the shift of the polymerase. So, this goes to the P site and the A site remains free. But depending on the codon the next tRNA will be brought.

This is catalyzed by ribozyme. Ribozyme catalyzes this reaction. If you have a pyrimidine base you have lot of basic nitrogens. It is not always the most nucleophilic or most basic nitrogen that is going to assist the catalysis it is the proximity effect. So, the way it is done is that this nitrogen is proximal to the see what happens here. So, that becomes a tetrahedral intermediate. This O minus tetrahedral intermediate is stabilized by the protonated form of the adenine base.

So, that will now stabilize. So, it now assists in generation of the tetrahedral intermediate which is required for the peptide synthesis. So, this base is provided by the RNA itself and this is called a ribozyme. The RNA acting as enzyme are called ribozyme. Ribozymes are also include the spliceosomes which takes care of the intron.

I think, we just have a quick glance of the recollection of whatever was taught. It is a complicated area i.e. replication transcription and translation. So, many enzymes, proteins are involved and so many mechanisms are there. So, I thought that before going to the next topic will have a revisit of those topics. So, that has been done now. So, in the next session we will go to our next topic that is the amplification of DNA.

Thank you.