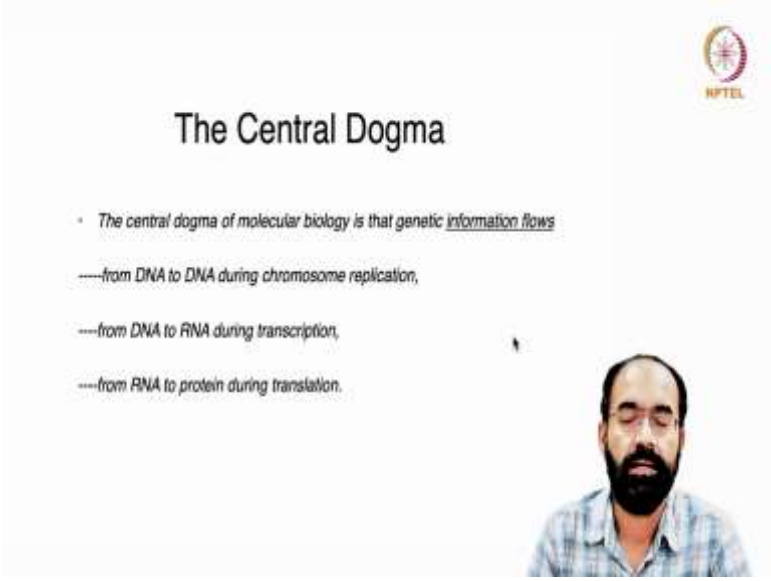
RNA Biology
Prof. Rajesh Ramachandran
Department of Biological Sciences
Indian Institute of Science Education and Research, Mohali

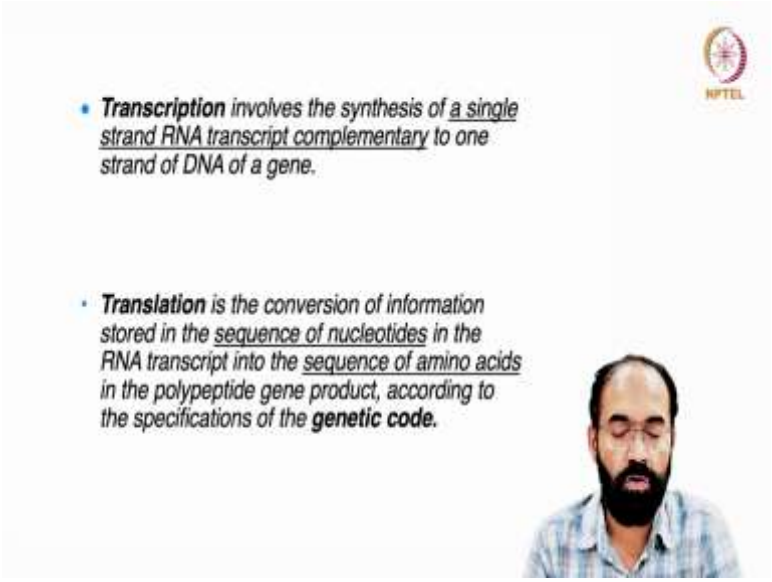
Lecture - 11
RNA Transcription: Initial Steps

(Refer Slide Time: 00:21)



Welcome back to another session of RNA Biology. So, we stopped here in the previous class that the central dogma which deals with making a copy of DNA from an existing copy and then making DNA as a template or using DNA as a template to make RNA and then RNA that is formed from the DNA should get converted to protein. So, this is what central dogma is dealing with; no matter whether it is a eukaryote or prokaryote or even viruses the same procedure has to be followed.

(Refer Slide Time: 01:04)



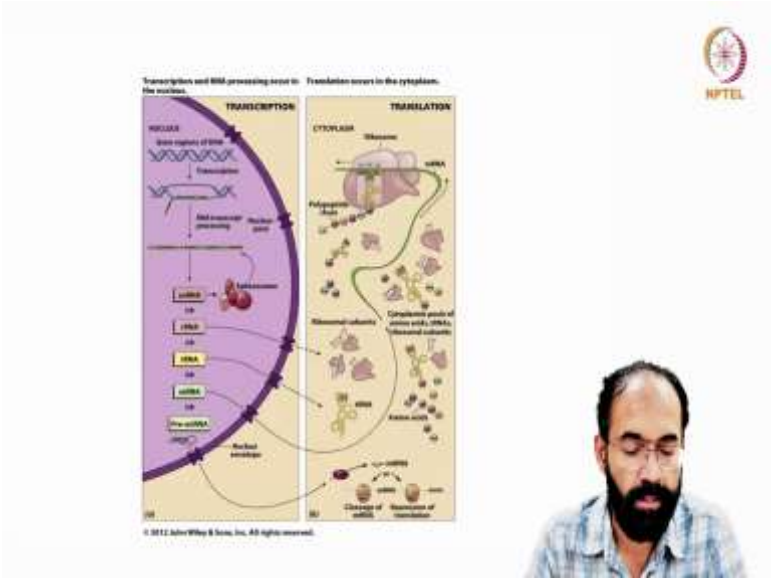
And, transcription involves a single stranded RNA transcript that is complementary to one strand of the DNA or a gene. Remember, when you say complementary strand, they are mirror images of each other. A pairs with T, G pairs with C in the case of DNA. So, the transcription involves the utilization of one of the strands in a DNA. In the DNA, we can classify these strands as a sense strand and anti-sense strand.

So, the sense strand basically codes for a mRNA actual mRNA sequence and it has information or codons that has to be translated into amino acid sequence in the cytoplasm. However, the other strand that is the antisense strand is acting as the template or in other words the sense strand is the non-template which is not the template, antisense strand is the template strand.

So, that the antisense strand when acting as a template, it can give rise to a exact copy of the sense strand in the form of RNA. So, this idea should be clear in your mind. And, this formed RNA undergoes translation and this is basically conversion of the information that is stored in the sequence in the form of sequence information or sequence of nucleotides in the RNA transcript into a sequence of amino acids.

So, every 3 base in the coding sequence of an RNA, now codes for 1 amino acid. So, if you have got 9 bases, it will code for 3 amino acids. So, this idea should be clear and we refer to this as the genetic code. So, every 3 bases is nothing but a triplet code or triplet codon.

(Refer Slide Time: 03:09)



And, the transcription takes place in the nucleus. So, in a eukaryote there is always a clear-cut nucleus that demarcate the genetic material from the rest of the environment that is the cytoplasm. So, in the nucleus you have a double stranded DNA and then one of the strands acting as a template strand.

And, this is a complementary strand or a nonsense strand to the top existing sense strand. So, this antisense strand acting as a template strand make a copy of the RNA with the help of DNA dependent RNA polymerase. And, now this RNA is identical to that of the sense strand or sense DNA strand and this RNA that is formed has to undergo process called RNA maturation and with the help of spliceosome.

And, this process is called mRNA, pre mRNA splicing and this is further enabled or made possible with the help of snRNA, small nuclear RNA. And, other RNA genes that are transcribed in a similar fashion are ribosomal RNA, transfer RNA, pre micro RNA etcetera. mRNA we have seen it that codes for proteins, we call them as a messenger RNA.

So, no matter which RNA you are talking about snRNA, rRNA, tRNA, mRNA or pre micro RNA, they all encoded in the DNA. It has to be produced from the genetic material itself. Once it is matured, it has to come out through the nuclear pore complex. So, nucleus is a three-dimensional structure which has got openings in multiple places and we refer to this as the nuclear pore.

Just like doors in your house, if you want to move from one house one room to another you have doors, to enter from out of the house into your house you need to have door. So, nuclear pore is quite stringent way, it in a systematic manner the nuclear pore complex functions to regulate the movement.

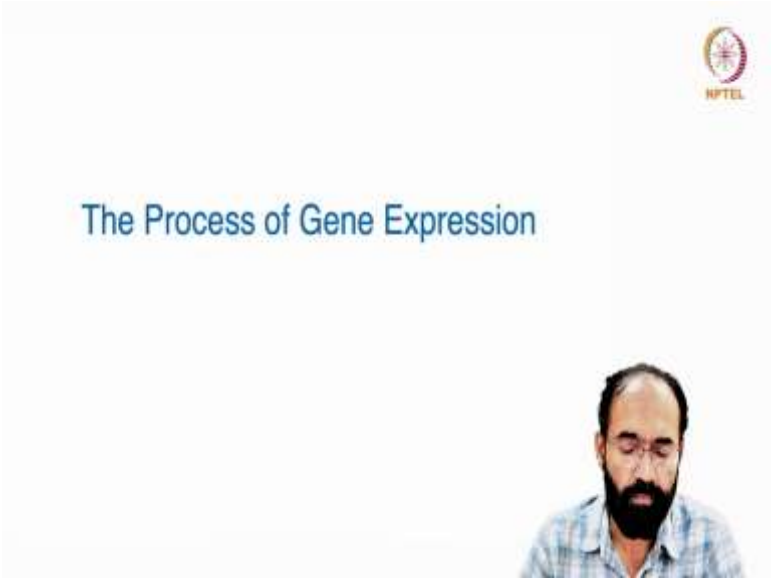
So, any mRNA present in the nucleus cannot go to the cytoplasm unless it qualifies or it pass the test of the nuclear pore complex. We will see about that more in detail in the subsequent classes. So, the cytoplasm now it is waiting for this mRNA and also the tRNA, rRNA etcetera.

rRNA and tRNA will perform in a different way, because rRNA has to give rise to the ribosomal complex and tRNA has to mature. And, we have seen it how tRNA undergoes maturation in the earlier classes and amino assimilation of the tRNA has to take place; so, that it can participate in the protein synthesis.

Now, these ribosomes will recognize the mRNA that is coming out of the nucleus and this can allow the protein translation to take place in a smooth manner. So, every 3 bases present in the mRNA codes for 1 amino acid and the polypeptide chain that comes out of this mRNA ribosome complex will find its way into the cytoplasm and the protein will have its own way of functioning.

So, this is further tweaked, this process of translation is further tweaked by other proteins such as micro-RNA and other non-coding RNA etcetera. We will see them one by one as we discuss those topics.

(Refer Slide Time: 07:09)



So, the process of gene expression is quite a lauded terminology. Gene expression is not mere production of an RNA or mere production of protein. It is basically production of a character, production of a feature from the stored genetic information; that means, it can be stopped in multiple levels rather the stopping at multiple levels is important for the survival or effective use of the information.

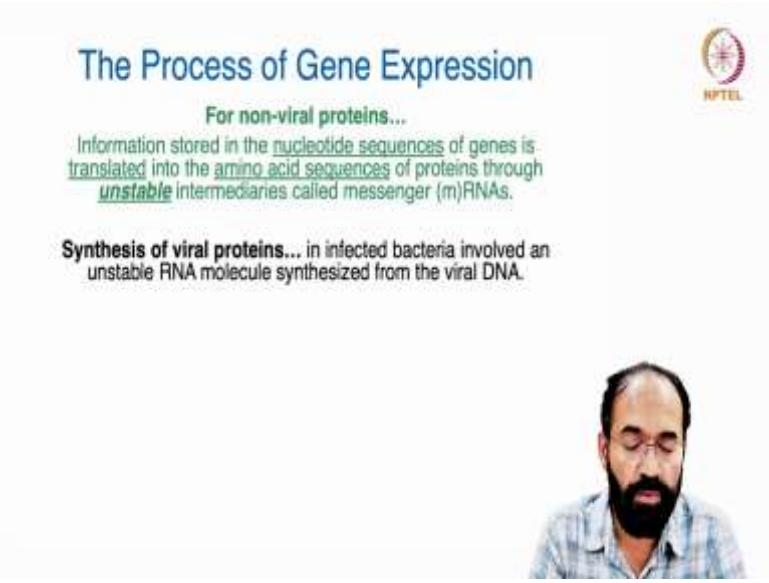
Sometimes, a misplaced information or an information that is unwanted can give you give lots of trouble to the cell or it can be problematic. To prevent this what the system will do if a protein is produced in an unwarranted manner, then what it will do? It will simply get rid of it or it simply do not allow this mRNA to get into protein and this mRNA will be marked for degradation.

So, this is just like you have a pen in your hand and you want to use it for writing, suddenly pen dropped from your hand. So, what you do? You will pick it up and restore normalcy. So, this is how you are able to make use of your pen. So, this is what you have done is regulation.

So, your hand helped in regulating the process of writing. If pen is not in your hand, you simply cannot write. Same logic applies whether an mRNA present more or less has to be regulated. If there is a less mRNA, it has to be produced more. If there is more mRNA, say you have got 6 pen in your hand; what will you do? You will get rid of 5 of them.

Or, if you are not wanting to write at all, then what will you do? You will get rid of all the 6 of the pen from your hand so, that you do not have to write. So, this way you regulate your process of writing, cell also do the same way. So, that is the process of gene expression is a stringently regulated one.

(Refer Slide Time: 09:08)



So, if you look further for some of the non-viral proteins, what you see? The information stored in the nucleotide sequence of genes is translated into amino acid sequence of proteins through an unstable intermediate called mRNA which is usual for bacteria, other eukaryotes and even viruses.

But viruses their genome is small. So, they have certain unique elements that can fool the system or they can mislead the system that is eukaryotic system as if it can recognize a middle portion of the mRNA as a tip of the mRNA or the 5 prime end of an mRNA. And, one such example are internal ribosome entry site. So, this allows polycistronic function.

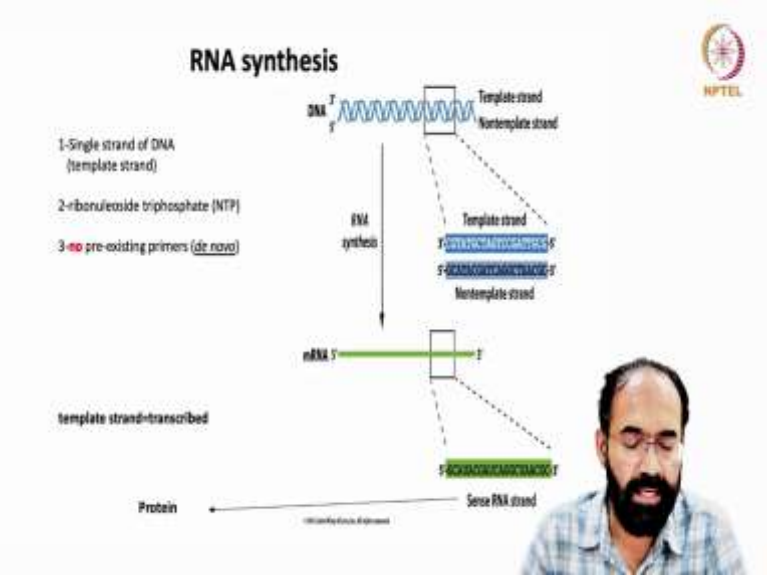
Normally, in eukaryotes the genes are monocistronic, monocistronic means 1 mRNA, 1 protein. But, in some case like such as bacteria, they have polycistronic; 1 mRNA can give rise to say if an m RNA is 100 base region; let us it is an example. 1 to 30 will code for 1 to 30 base will code for 1 pro amino acid sequence or a protein and say 35 to 55 will code for another protein and say 65 to 100 will code for another protein.

All part of one mRNA, but it does not happen in the case of eukaryotes. Similarly, viruses can fool the system, but remember the virus has to virus mRNA has to function in a eukaryotic system. So, eukaryotes are not at tuned to produce more proteins from 1 mRNA. So, the viral mRNA contains certain unique sequences that can fool the ribosomal machinery of the eukaryotes.

So, that 1 mRNA produced from the virus can give rise to multiple proteins. And, this way they enhance the chance of their survival with minimum genetic information. So, synthesis of viral proteins in an infected bacteria, when a bacteria is infected with virus, the bacteria eventually dies. So, we refer this as a bacteriolysis and the virus that infects bacteria are called bacteriophage.

One good example is the lambda phage. Many t47, many different types of phages are there; lambda phage is one of the well classified, well studied example of virus infecting the bacteria. And, they when a bacteria get infected with a virus, they involved an unstable RNA molecule that is synthesized from the DNA and this is allowing the propagation of the bacteriophage in the host. Same way a eukaryotic cell infected virus also undergoes the influence by the virus.

(Refer Slide Time: 12:13)



So, if you look into here the RNA synthesis, a single strand of the DNA template strand can act as the source information; that means, a template that can make a copy of the opposite strand or that is making a complementary strand that is equivalent to opposite

strand or the sense strand. So, a single strand of the DNA takes place or participate in the copying of the RNA.

And, this process is called transcription and the monomer used are nucleoside triphosphate or NTP. And, if a DNA is being produced, we call it as deoxy ribo nucleoside triphosphate and the there is no requirement of a primer. For a DNA synthesis to take place, you need to have an RNA primer which is a indicator of an RNA world hypothesis.

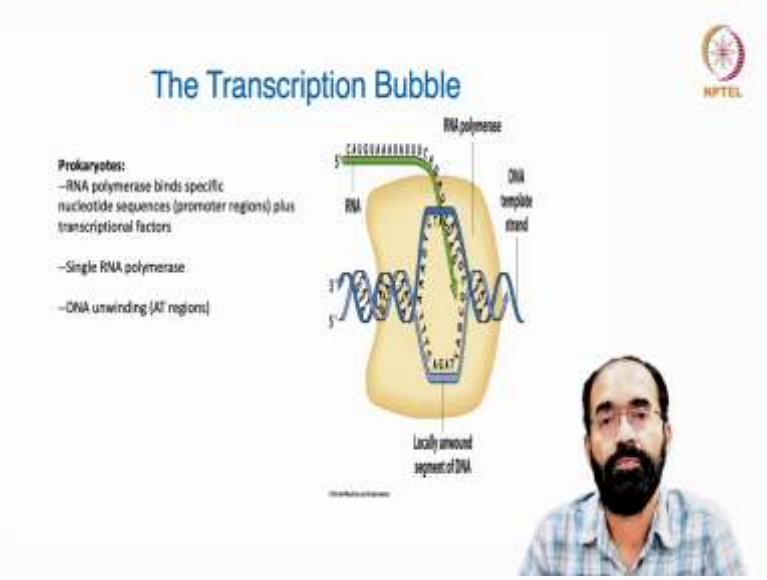
We have discussed about RNA world hypothesis. For a DNA to start replication, every time you need to have a small RNA primer to initiate the DNA replication. But, for the transcription that is not required, you can straight away start the transcription if you have an enzyme, RNA polymerase is present.

And, the template strand is the one which helps in the transcription, that gets transcribed. The non-template strand which is basically the sense strand do not participate in the transcription. So, this is a DNA double strand and then you have certain region, that is having unique sequence that allow the RNA synthesis to take place.

And, this mRNA have certain features that allow the maturation of the mRNA. Sometimes, if mRNA transcription is continuing, it need to be stopped even if the transcriptional machinery is continuing further it need to be cleaved. And, certain sequence features in the mRNA, the newly formed mRNA helps in cutting it at specific loci. So, that is what we should understand.

This mRNA, unified or well-defined mRNA is now undergoing protein translation and you end up getting a protein product coming from the RNA that is formed.

(Refer Slide Time: 14:39)



And, now we will learn more about the transcription bubble. As the name itself indicates, the transcription RNA transcription takes place in the form of a bubble both in prokaryotes and eukaryotes. So, the enzyme responsible for this transcription is called RNA polymerase that binds to specific loci, specific location or specific nucleotide sequence.

We call them as promoter regions or regulatory regions. We will see more in detail about the gene regulations, how the various protein factors are contributing to the gene regulation. But time being we are understanding the basics of transcription in a prokaryotes.

And, a single RNA polymerase is important; that means, in eukaryotes it may not be single that is why in prokaryotes we use the word single RNA polymers. We will see that later, about the eukaryotic transcription. And, DNA unwinding has to take place and usually they are AT rich region. AT means what? Adenine and Thymine.

Means, we know adenine is pairing with thymine and if thymine is present adjacent, then it will pair with adenine. So, basically it is a AT AT AT or sometimes AAT or ATT like that. So, basically AT region means, does not really mean that it has to repeat like AT AT AT, not like that, it can be predominantly AT rich. And, we know A pairs with T is a double bonding which is weaker compared to a G pairing with C which is a triple bond.

So, this unwinding is much much easier when you are talking about AT rich regions. Nonetheless, this area has to be contributing to the melting or the unwinding or opening of this area, but it is reasonably strong. So, do not think that AT rich regions are weak, AT rich regions are less connected or loosely connected to the opposite strand the opening becomes easy, that is how the evolution has made the regulatory regions.

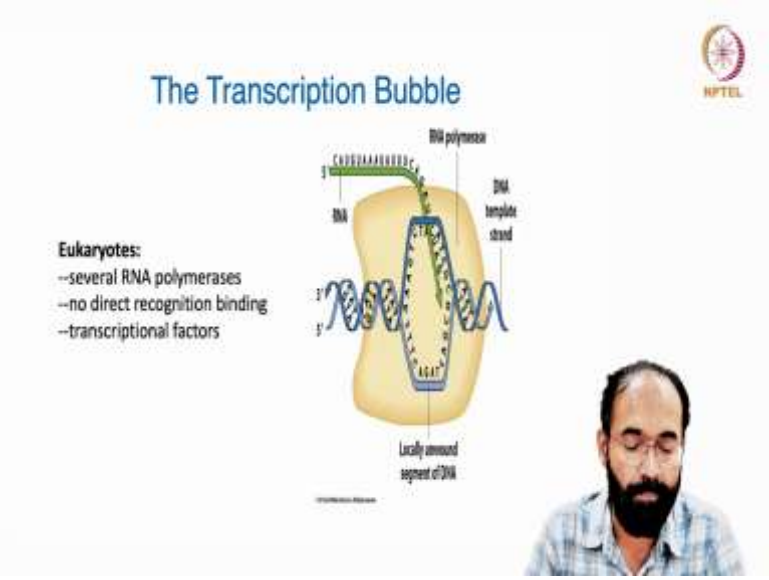
So, that the RNA polymerase can make a copy of RNA from one of the template strand. As you can see in this picture, there is a DNA double helix and downstream also there is a DNA double helix. But locally there is an opening and one of the strands is acting as a template and the locally unwound DNA now keep migrating.

Because, this yellow color circle or this shape what you are seeing is the RNA polymerase itself and this RNA polymerase allows local unwinding of the DNA. So, that the double helical structure is locally damaged and this continues like a bubble. So, if you have a soap bubble in water that bubble can move from one place to another right.

Same way this bubble moves on the DNA, while opening the DNA slowly and slowly and the upward and downward regions of the bubble, the double helix is restored back. So, it is opening very gently. Like if you have a rope, plastic rope in your hand which is tied down.

So, if you twist in the opposite direction, the strands of the rope slowly open up right. Same logic applies, you can open up the DNA and this is what we refer to as transcription bubble.

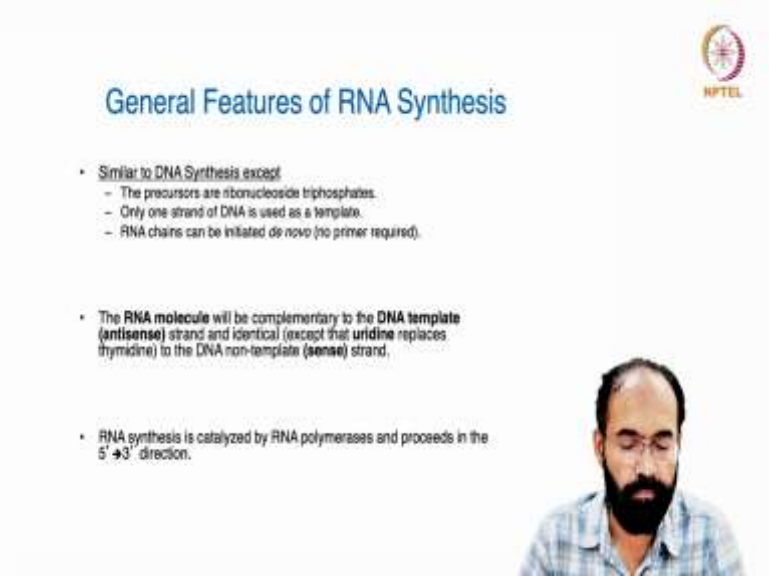
(Refer Slide Time: 18:11)



And, in eukaryotes process is same except that several RNA polymerases are there. RNA polymerase is a complex of multiple protein and there is no direct recognition onto the binding takes place, because it is a collective effort. No direct recognition is there on to the binding.

And, also several transcription factors are there. We can call them as general transcription factors or specific transcription factors; both comes on to action for a eukaryotic transcription. But the transcription bubble remains the same, but the number of protein participate is more than that of prokaryotes.

(Refer Slide Time: 19:02)



Now, if you look the general features of RNA synthesis, it is very much similar to that of DNA synthesis; that means, in DNA synthesis both the strands act as the template. It has to make a copy of itself. So, the precursors are ribonucleoside triphosphate, unlike deoxy ribonucleoside triphosphate in the case of DNA synthesis, because here we are producing RNA. Only one strand of DNA is used as a template, another difference.

And, RNA chains can be initiated de novo; that means, no primer is required because genes can be located in different parts of the genome. You cannot expect an RNA primer corresponding to that sequence be available every now and then. So, system has evolved in such a way that RNA polymerase can start spontaneously at selected locations. Spontaneously means not at random location, selected location which has the so, called promoter sequence elements.

So, the RNA molecule will be complementary to the DNA template that is the antisense strand and identical except for the uridine replace the thymine. It is identical to the DNA which is non-template, but that is a sense strand. So, this should be very clear. The concept of template and non-template sense and anti-sense, this idea should be very clear.

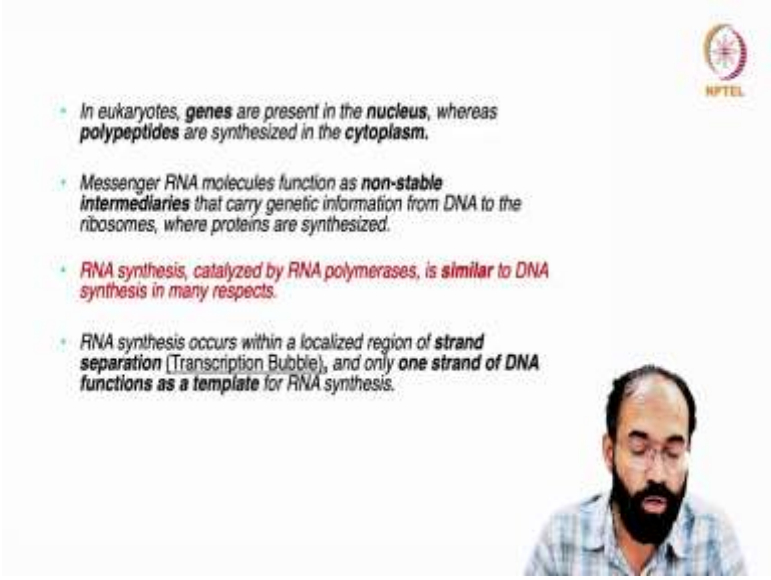
And, once you have the concept clear, you will not have confusion. So, the RNA synthesis is catalyzed by an RNA polymerase and it proceeds only in one direction, 5 prime to 3 prime. Like you know the pentose sugar, the ribose sugar have got 5 carbon

atoms and carbon number 3 and carbon number 5 are important in forming the phospho diester backbone.

The 5 prime always will have a phosphate group and the 3 prime will always will have a hydroxyl group as the net result. So, the when 2 nucleotides are coming together, the 5 prime phosphate participate in a phosphodiester backbone formation with the 3 prime OH onto another nucleotide ribose sugar. So, end result will be you have a 5 prime phosphate group projecting out and extreme end of the nucleotide, you have a 3 prime OH group projecting out.

So, the nucleotides always or the polymer RNA always have a directionality, 5 prime to 3 prime direction. So, the synthesis of RNA takes place in one direction or unidirection, that is 5 prime to 3 prime direction. And, this rule cannot be reversed, you cannot have an RNA produced from 3 prime to 5 prime direction.

(Refer Slide Time: 21:54)



In eukaryotes what happens? The genes are present in the nucleus whereas, the polypeptides are synthesized in the cytoplasm. And, the messenger RNA molecules function as a non-stable intermediaries that carry the genetic information from DNA to the ribosome where the proteins are synthesized. When you say non-stable means, it is stable, but it is not permanently present. So, that is what we should understand.

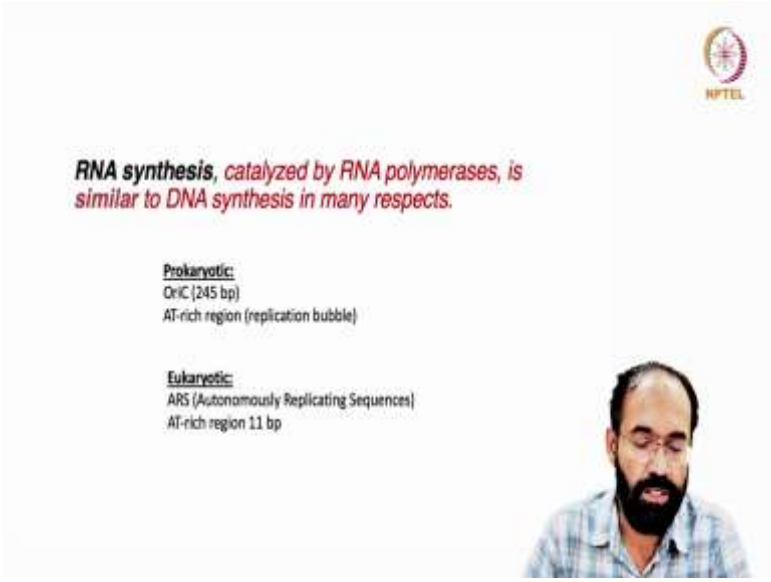
And, during RNA synthesis, it is which is catalyzed by RNA polymerase is very much same or similar to the DNA synthesis in many respect. The only difference is uracil comes in and instead of deoxy ribosugar, the ribose sugar containing NTPs come in. And, the RNA synthesis occurs within a localized region of strand separation; that means, the DNA double strand has to open up. It happens in a localized region, not the entire DNA opens up.

It opens up only in a place where it is actually required and this is called transcription bubble. It literally look like a bubble, when you have the you know RNA polymerase etcetera coming into picture. And, only one strand of the DNA functions as a template for the RNA synthesis. Both strands can participate, but not in the same region.

Say, if region number 1 to 1000 base maybe top strand is participating, but maybe from 2000 to 5000, the bottom strand may participate as a it is a different gene. So, do not think that in a DNA only one strand contains the gene ok, one strand only contains the gene in that particular context or the meaningful or coding information is present on one strand in that particular context.

Whereas, after some bases or some thousands of base pairs down the line or upward 5 prime direction or 3 prime direction, you can have the other strand coding information for proteins. So, there is no rule that only one strand will act as the template. In a given scenario, a given context that only is going to a given strand act as the template. But, for another gene the opposite strand may be containing the information.

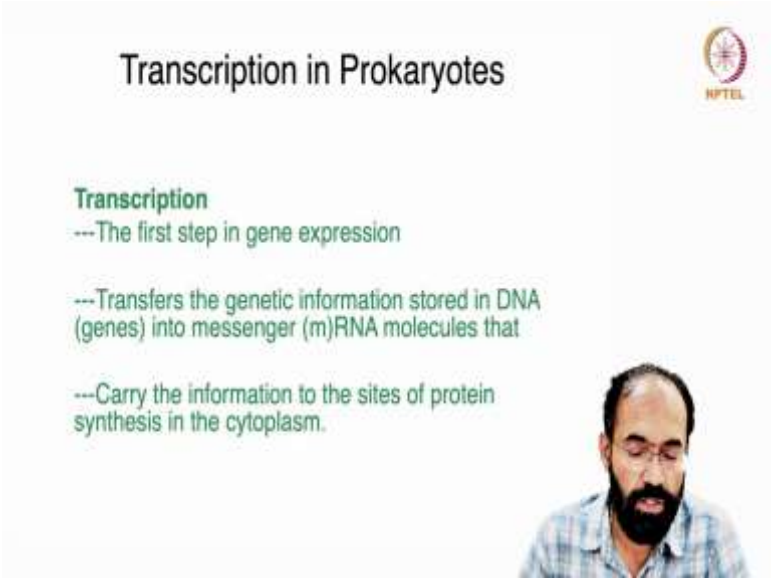
(Refer Slide Time: 24:25)



So, RNA synthesis is catalyzed by RNA polymerase is very much similar in many respect whereas, in prokaryotes you have a region called OriC which is basically have around 245 bases in length which is an AT rich region that is basically initiating the DNA, initiating the DNA unwinding or the formation of the replication; that means, making a copy of the RNA from one of the strand.

Whereas, eukaryotes you have certain region called ARS which is Autonomously Replicating Sequences which is similar to the transcription or the RNA production. But these regions are rich in AT and this rich in AT sequences allow the opening much much easy and this is very much similar to that of DNA synthesis.

(Refer Slide Time: 25:28)



So, if you look further into the transcription in prokaryotes, transcription is the first step in the gene expression. And, it transfers the genetic information stored in the DNA to genes into messenger RNA and these are the molecules that eventually carry information for the protein coding purpose. So, we will address this transcription in prokaryotes more in detail in the subsequent classes.

Meanwhile, you should clearly focus on what are the topics that is covered and we cannot postpone for in the end to sit and watch. So, every day if you are looking every section very focus fully, subsequent classes will become easier for; you to follow you do not have to refer back on a previous video. So, I would strongly request that you follow it up systematically and also refer the textbook assignment required.

Thank you.