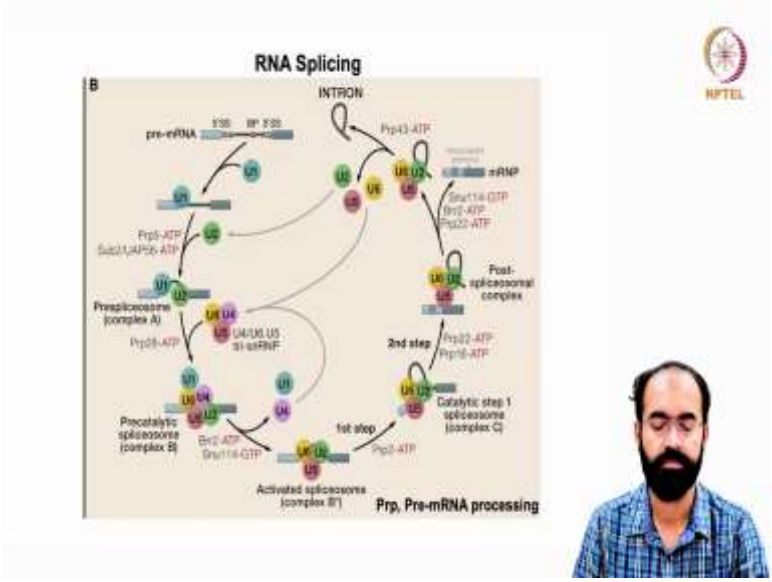
RNA Biology
Prof. Rajesh Ramachandran
Department of Biological Sciences
Indian Institute of Science Education and Research, Mohali

Lecture - 27
RNA Splicing, Export and Stability: SMN Complex

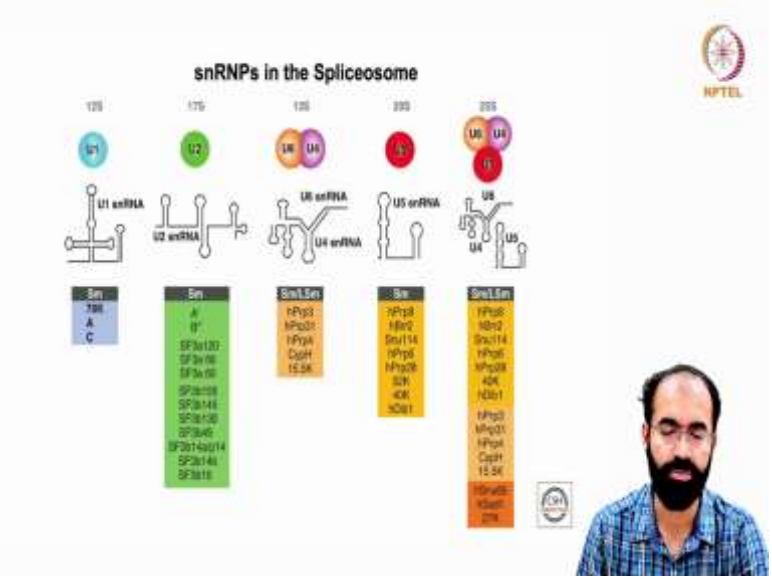
(Refer Slide Time: 00:21)



Hello everyone, welcome back to another session on RNA Biology. So, we were here in the previous class that is RNA Splicing in detail. And we were seeing what are the proteins involved, what are the complexes that are formed, a complex A, complex B, complex B star and complex C and the post spliceosomal complex. During which there are 5 snRNPs are utilized U1, U2, U6, U4, U5, they are utilized and along with the utilization of 10 ATP molecule and 2 GTP molecule.

So, splicing is a tremendous task with lot of energy consumption. So, that is why in a needless to say the production of RNA also require lots of energy that is ATP yielding reaction. That is why the system regulates these snRNA their production, their function, their interaction, their stability, their usability, etcetera in a stringent manner. Let us now look in detail how individual snRNAs form complex and what is their structure.

(Refer Slide Time: 01:45)



So, here you can see U1, U2, U6, U4 pair, U5 and U6, U4, U5 pair. So, usually they can, these are all the different stages in which the complex during the single splicing event, this many complexes form. U1 has around 12S Svedberg unit is its size and this is the size and the shape of secondary structure of U1 snRNA and it has got some Sm protein interaction.

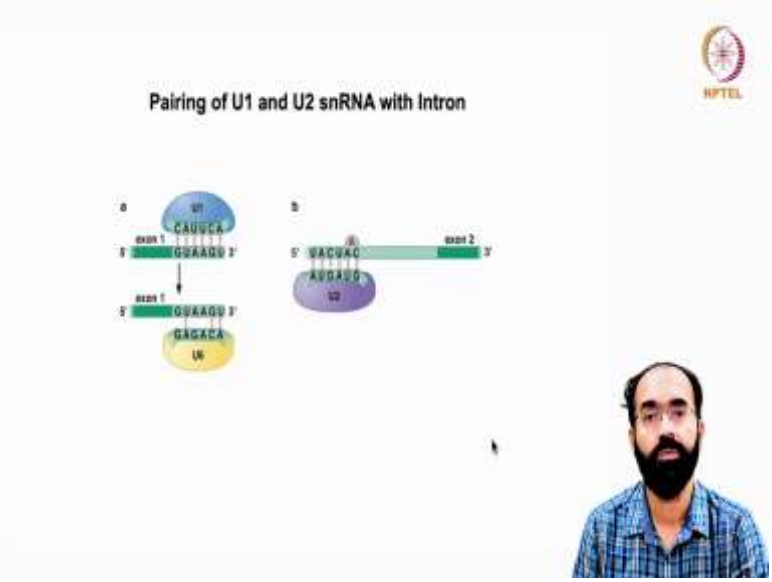
We will see about Sm proteins in detail in the future class. Because now if I start talking about Sm proteins it will be completely you know out of context for time being. So, time being you understand that each of this snRNA this is their structure, this is the U2 snRNA 17S because picture itself you can make out, this is 12, this is 17 and the size also is much high for the U2 snRNA.

And these are all the possible complex like A and C complex and A star and B star complex it is formed and you have multiple proteins associated with this U2 snRNA. And you have the 13S, U6, U4 complex as you can see here this is two RNA basically paired together, U6 snRNA and U4 snRNA.

They pair and of course, they flip also they are pairing and we will see that more in detail in some of the subsequent slides. And they also have a bunch of proteins interacted. Their names are not important, but this slide is basically to give a clear idea that how complex this snRNPs are.

U5 snRNA have got around 20S size and it interacts with this much bunch of protein. And this complex U6, U4, U5 which is around 25S molecular weight and this is their interaction pattern and you have bunch of protein interacting with them as well.

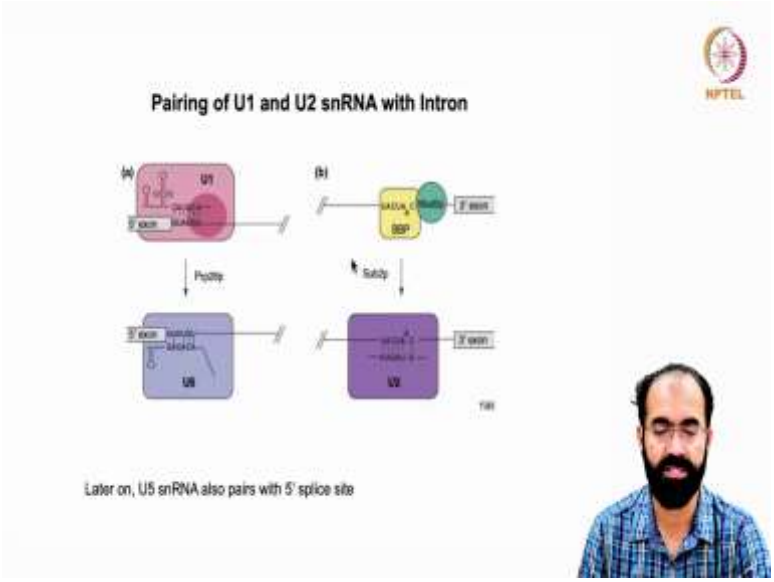
(Refer Slide Time: 03:57)



Now, let us see how the pairing is taking place, the U1 and U2 snRNA, how it is assembling in the intron. We all know by now that U1 is the first one to join the complex and it binds on to the exon intron boundary simply by sequence pairing. And this is the exon and this is the intron part and it is pairing very nicely.

And it can further be influenced by U6 snRNA in that place not now later stage. And at branch point you have the pairing of the U2 that is in the intronic sequence. So, this is in the exon intron boundary and this is purely in the intronic sequence and the pairing is taking place.

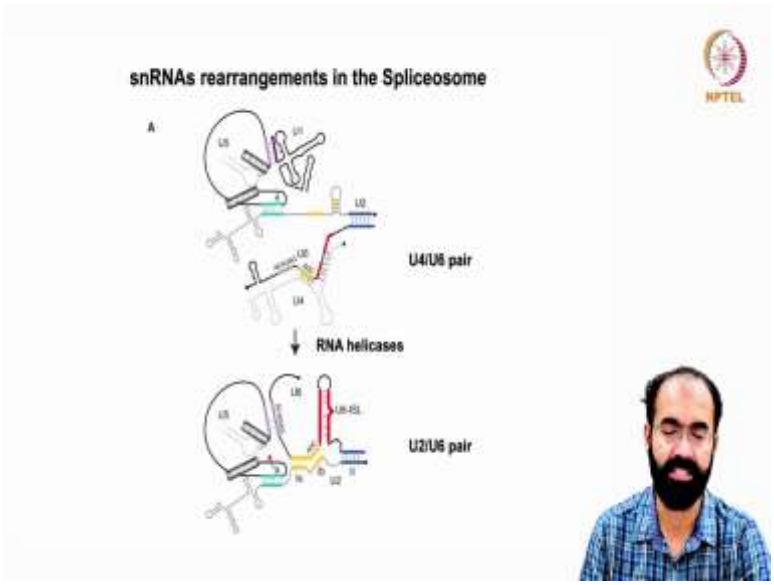
(Refer Slide Time: 04:49)



Let us see what happens next. So, U1 snRNA paired to the exon intron boundary and this eventually will be occupied by U6 at later stage once U1 leaves the complex. At the branch points you have the U2 snRNA that joins and this is with the help of a protein Sub2p it gets into a position such a way that the branch point A is projected out it is no more welcome no more participating in the pairing even with the U2 snRNA.

So, that this A can easily participate in the transesterification reaction. Now, let us see how the actual splicing takes place.

(Refer Slide Time: 05:35)



So, snRNA rearrangement takes place during the splicing event. So, we have to see very clearly that U1 snRNA has paired and also look that this is the one exon and then you have the intron and then continuing with the exon here. So, U1 snRNA paired a portion of the exon intron boundary and this intron is going huge and it has a branch point where the U2 snRNA is pairing and the this close to the second exon. The first exon intron branch point and the second exon.

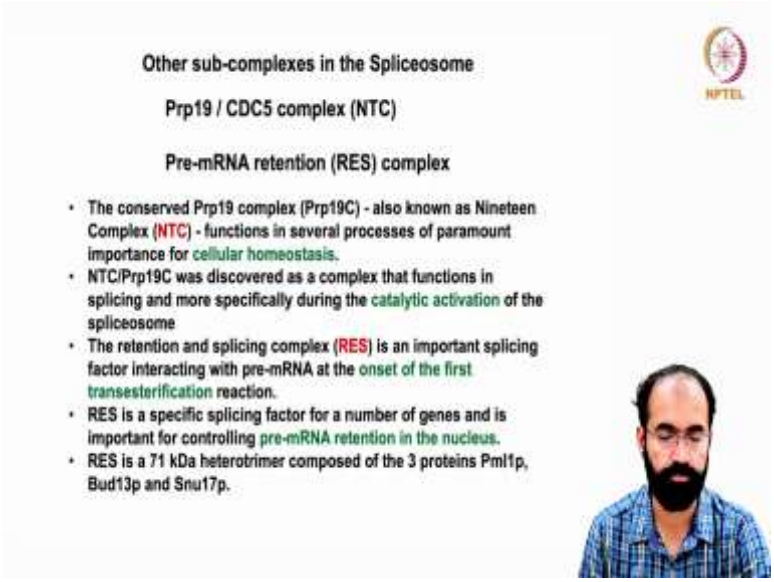
Now, U1 has paired and U2 also has paired. But remember the U2 has got a much longer piece that is projected out which is welcome to occupy a small portion of U6 and U6 is not free. U6 now comes paired already with U4. So, U4 U6 pair now pairing partly with the U2. And then what happens?

There is an RNA helicase activity taking place and because of this RNA helicase activity you see that the U1 which had paired here is now replaced by U6 which was very much here. And the U6 came here and when U6 comes it is bringing the U2 also because U2 is in the branch point.

So, now U1 is replaced or displaced by the U6 and when U6 comes it brings in A the U2 also because it is already paired with the U2 along with the U4 U6 pair. So, the purpose of U6 occupying the U1 site is to bring in the U2 or ideally the branch point closer to the first exon. So, that the 5 prime splicing can happen or the first transesterification can happen.

So, now you have U2 U6 pairing. Because initially it is started with the U4 U6, now you have U2 U6 pairing. And now U1 is displaced and U2 came closer to the 5 prime splice site. So, that the branch point can access the 5 prime splice site in the exon intron boundary.

(Refer Slide Time: 08:12)



There are other sub complexes also in the spliceosome and they are called Prp19 and CDC5 complex which is also known as NTC complex and then Pre-mRNA retention complex it is called RES complex.

So, the conserved Prp19 complex it is also known as Prp19C is also known as 19 complex that is why NTC come from 19 complex. And it functions in several process of extreme importance in the cellular homeostasis of any given cell. Because these complexes regulate the rate and the propagation of the rate of propagation of the splicing.

So, the NTC Prp19C was discovered as a complex that functions in the splicing and more specifically during the catalytic activation of the spliceosome. Because we know many splicing signals are present in the snRNAs are there snRNPs are there and the splicing location and the signals are present in the pre mRNA etcetera.

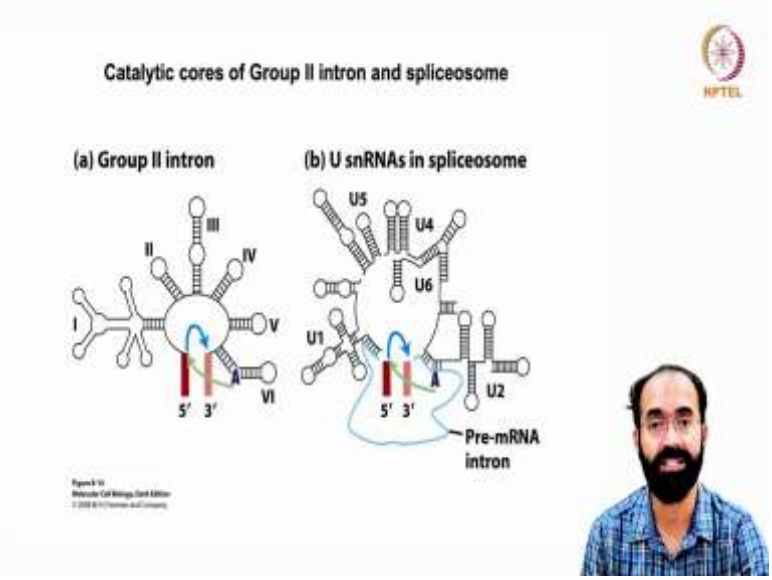
But something has to kick start. So, that is why this complex becomes important to energize them. That is like if you have a car, it has an engine, it has got fuel everything car will not run you have to kick start, you have to start the ignition, you have a spark plug and you start the ignition.

So, the retention and the splicing complex that is called RES complex it is an important splicing factor interacting with the pre mRNA at the onset of first transesterification reaction. Because without which it will not proceed in a spontaneous manner.

So, RES is a specific splicing factor for a number of genes and is important for controlling pre mRNA retention in the nucleus. So, RES complex also make sure that no unwanted or unprocessed RNA move to the cytoplasm. So, it retains the incomplete or intermediate or half complete RNA pre-mRNAs are retained until it is ready for export to the nucleus.

So, RES is a 71 kilodalton heterotrimer composed of 3 proteins and they are Pml1p, Bud13p and Snu17p. So, these 3 together constitute the RES complex. Why this is mentioned here now? Because we should not assume if U1, U2, U3, U4, U5 these are there they cannot continue to do the splicing. There need to have a kick starter or a regulator.

(Refer Slide Time: 11:01)



So, the catalytic core of the group II intron and spliceosome have lot in common. Group II intron we have seen it that is it is very much similar to that of the normal splicing.

It has the A at the branch point is utilized and it is a ribozyme. So, if you look group II intron the secondary structure mimics that of a typical spliceosome. Like you can see here there is one exon another exon and you have the arm 1, 2, 3, 4, 5, 6 same way you have multiple arms which is more or less the same when it comes to the splicing.

But here the entire sequence, entire arms, entire stem loops they are all formed by the group II intron sequence itself. Whereas, the these structures are formed in a typical

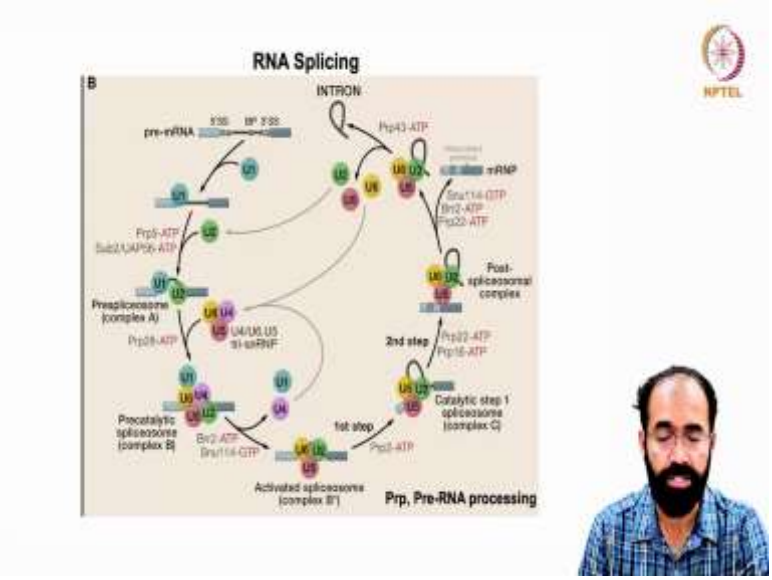
snRNA mediated splicing it is formed with the help of various snRNA such as U1, U5, U4, U6, U2 etcetera.

So, these are external RNA molecule that contribute to the formation of a structure that is similar to that of the group II intron alone. So, you may remember similar comparison between ribozyme and tRNA earlier one of the earlier class. So, many a times the splicing event require similar type of secondary structure formation, it cannot be done out of the blue the shape has to be in a particular manner.

This is just like every organism or every say if you if you are talking about a projectile motion if you are throwing whether you throw a 1 kilo subject or you throw a 100 kilo material or you throw a 1 gram material their action of gravity on them is same and if you throw it will form a parabolic shape, you throw it at 45 degree it will go reach a pinnacle and then gravity started bringing it down and based the velocity will come down and it will make a arc shape.

Same way here this kind of arc shape formation is very important. No matter whether you are talking about a self splicing intron such as group II intron or an assisted splicing such as snRNA mediated. So, this has to be very clear in your mind.

(Refer Slide Time: 13:34)

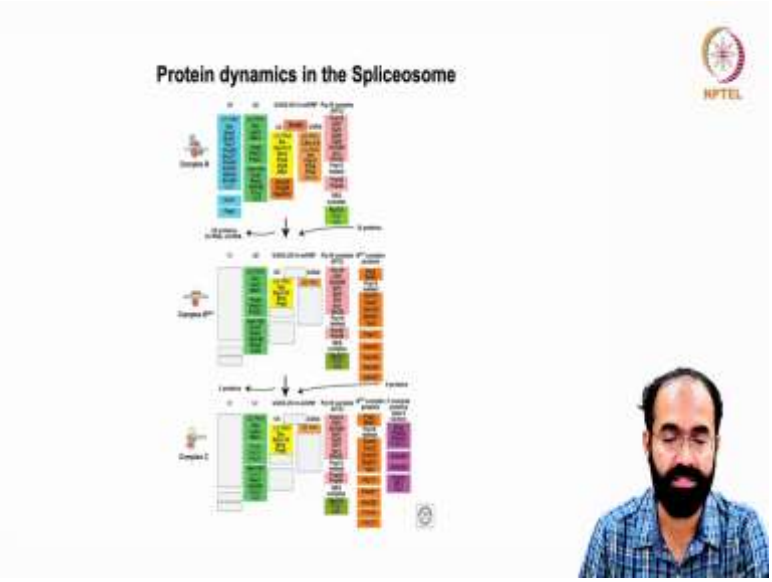


So, this we have seen it already for our convenience I have put it in once again to understand the splicing event that U1 joints U2 joints and then you have U6, U4, U5

joints. And then U1 and U4 release and the complex B star have got U6, U2 and U5. Because U1 is replaced and U6 occupied that place of the U1 as you saw earlier here.

It is occupied by U6 and then it will be further cleaved by various proteins that is in the second transesterification reaction and you have U6, U2, U5 released along with the intron that is removed from this pre-mRNA. And now they are further broken down. So, that the intron and this U1 snRNPs are released. So, this is the whole cycle.

(Refer Slide Time: 14:35)



Let us see more about the protein dynamics in the spliceosome. This may look little complex slide. So, do not worry about it. The purpose of showing this is to show how many proteins are associated with different snRNA at various stages of splicing that is complex B you have U1, U2 and also you have various other proteins that U6, U5 snRNP, how many proteins each of them what is shown in this box they are individual proteins they are associated.

Some of them are temporary association some of them are permanent association. And in complex B in an activated shape, it has got U2 mainly U1 is missing, U2 and then you have U4, U6, U5 pairing and some of the empty boxes indicate those proteins are missing. Because in this B star activated state you do not need those proteins.

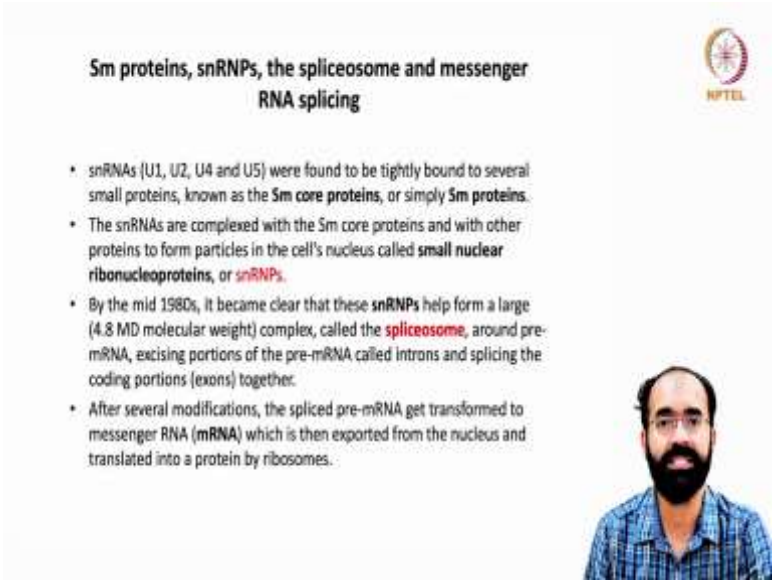
And you can also see in complex C stage U1 is missing and you have U2. So, and also you have this U6 and some of this NTC and RES complex proteins also contributing to

the splicing. The empty boxes indicates those proteins are not there in this particular complex. So, the purpose of showing this slide is for you to have a clear idea that any defect in the association or assemblage of any of this protein can have huge impact on the effective splicing.

So, that is why many a times in old age etcetera. If any of these proteins are in short supply this splicing even can get affected and the body will start suffering because of lack of proper splicing and the crucial genes. If they are affected then it can have deleterious effect on the organism survival itself.

But remember the defect may not be a permanent feature. Because the splicing is a temporary event. So, if one time it has caused a damage it made a defective protein then if it is fixed in the next round then the organism may not suffer that much. So, we also saw some splicing defect related genes in humans in a earlier classes.

(Refer Slide Time: 16:46)



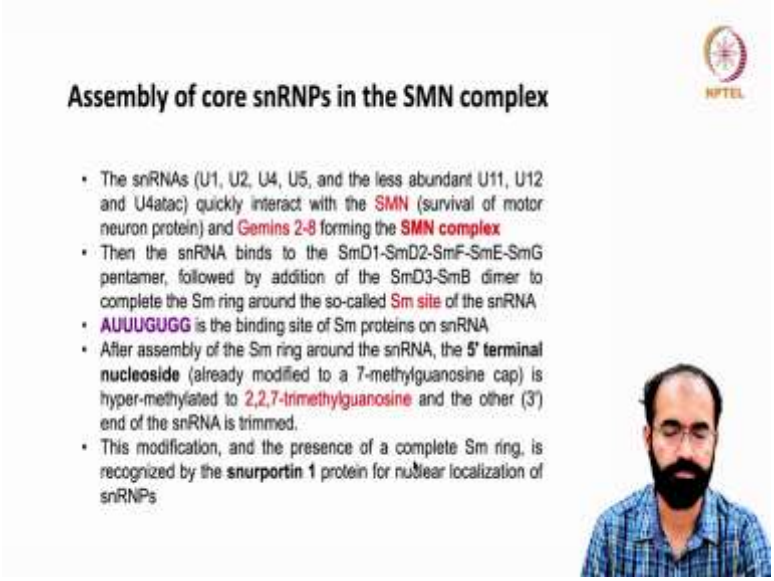
So, now let us see what is the involvement of Sm proteins and snRNP and also their contribution in the assemblage of spliceosome in the pre-mRNA. So, the snRNA we all know U1, U2, U4, U5 are the major ones and are found to be tightly bound to several small proteins and they are called Sm core proteins or simply they are called Sm proteins.

So, the snRNA are complexed with the Sm core proteins and with other proteins to form particles in the cell and especially in the nucleus and they are called small nuclear ribonucleoproteins or snRNPs. So, snRNP the terminology is formed because of the assemblage of snRNA with the Sm proteins mainly.

And by around mid 1980s it became very clear to the scientific community that the snRNP helped in a large complex that is roughly around 4.8 megadalton molecular weight complex and this is called spliceosome. And this occurs associated with the pre-mRNA and the excising it is helpful in excising portions of pre-mRNA which is nothing but the introns.

And this splicing of the coding sequence that is the joining of the coding sequences exons together. After several modifications the splice the pre-mRNA gets transformed to the mRNA that is ready for export into the cytoplasm which is then functioning in the cytoplasm with the help of ribosomes.

(Refer Slide Time: 18:35)



So, if you look further into the assembly of core snRNPs and also the formation of the SMN complex, the involvement of SMN complex. Let us see how it is important. So, snRNA we know clearly U1, U2, U4, U5 and there are two more we said the minor spliceosome U11, U12.

So, they are identical except that U1 is no more there instead U11 is there. And U2 is no more there whereas, U12 is there and which is used by a small group or small subset of pre-RNA or immature RNA for their splicing and also it make use of instead of U4 it makes use of U4atac snRNA. So, they quickly interact with the SMN and SMN stands for survival of motor neuron it is a proteins name and also with Gemins, Gemins 2 to 8 Gemin 2, Gemin 3, Gemin 4 like that 2 to 8 forming the actual SMN complex.

So, SMN is a important protein its contributing in various other cellular functions also it is important in the splicing as well. The snRNA binds to the SmD1-SmD2-SmF-SmE-SmG pentamer like it forms a circle followed by the addition of SmD3 and SmB dimer and to complete the Sm ring.

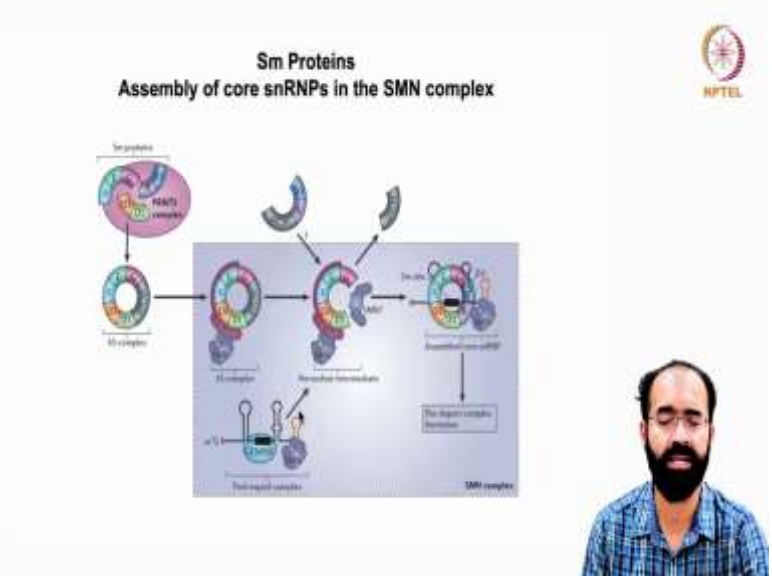
Sm literally appears like a ring that is why it is called Sm ring around the so called Sm site that is present on the snRNA to start with. That is the initiation of the formation of the snRNP. And the sequence of this binding is AUUGUGG is the binding site for the Sm proteins on snRNA.

This is quite interesting because normally RNAs are recognized by secondary structure, but here it is binding in a sequence specific manner just like a restriction endonuclease. So, after the assembly of this Sm ring around the snRNA the 5 prime terminal nucleoside that is already modified by 7-methylguanosine cap now is hyper methylated to become 2,2,7-trimethylguanosine and the other 3 prime end of the snRNA remains trimmed.

This happens in the cytoplasm. We know snRNA gets modified 1 cap or 1, 7-methylguanosine goes to the cytoplasm undergoes this modification with the Sm proteins and then gets hyper methylated come back to the nucleus. This modification and the presence of the complete Sm ring is recognized by the snurportin 1. This is the protein that is present in the nuclear localization of snRNP.

So, snurportin present snurportin 1 protein present in the nuclear pore complex recognize this hyper methylated snRNP and recruit back into the nucleus.

(Refer Slide Time: 22:06)



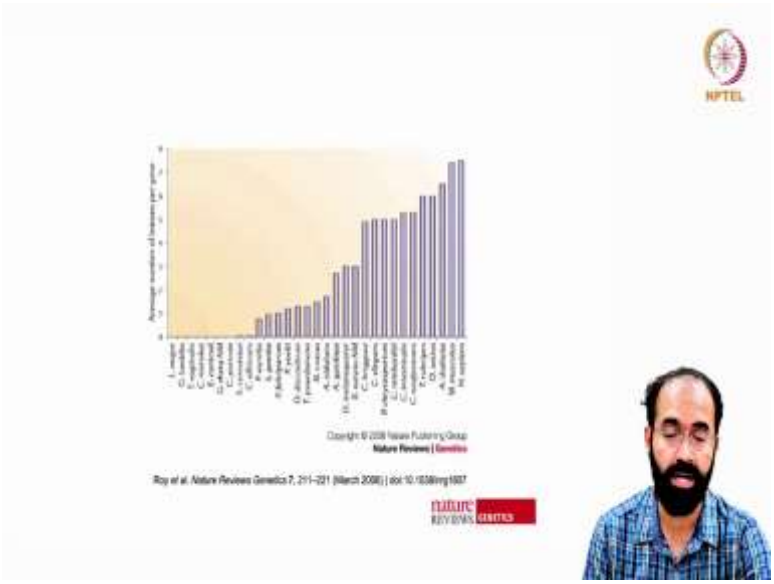
The same thing you can see in a pictorial manner here Sm protein and assembly of snRNP and SM SMN complex. So, Sm proteins assemble one by one and it forms a complete ring by the fusion of all this protein and what happens they assemble one by one along with the SMN Gemin 2 and also the Sm ring and it will continue to recruit more protein and eventually it forms a snRNP in combination with the snRNAs which is the pre import complex.

Once the hyper methylation is done it will go into the nuclear import takes place. So, formation of snRNP is a complex process. And it is happening in a sequential and stepwise manner. If this fails in some way or the other the splicing gets compromised and the compromised splicing can give serious problems with the functioning of a any given gene.

(Refer Slide Time: 23:12)

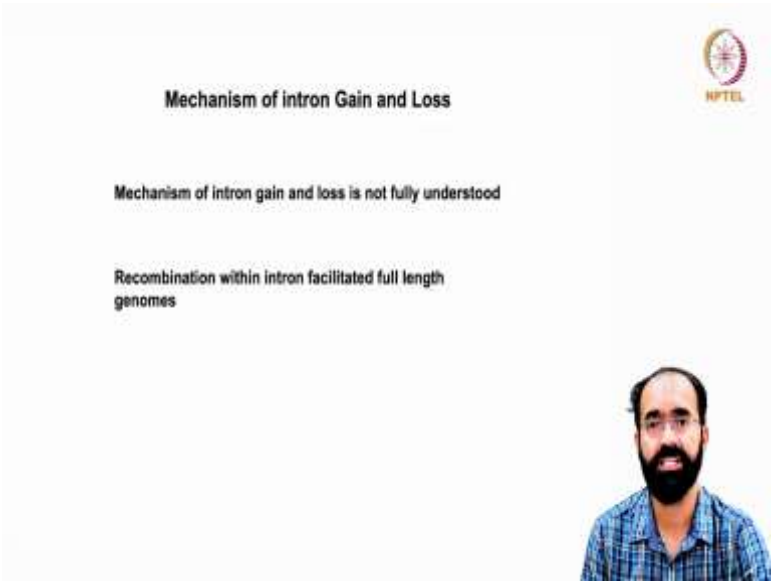
So, now let us address the mechanism of introns gain and loss. So, far we have seen evolution of introns, evolution maintaining, gaining etcetera. So, let us see what are the ways in which introns can increase in a gene or in an organism or we know that humans have an average 7 to 8 introns per gene. So, how can you increase or decrease the introns, how you gain or lose the introns is something important that we need to address?

(Refer Slide Time: 23:48)



We know this slide we have seen it earlier also as the complexity of an organism is increasing you have got more and more number of intron. So, around humans have got around 7 to 8 introns per gene same with mice also.

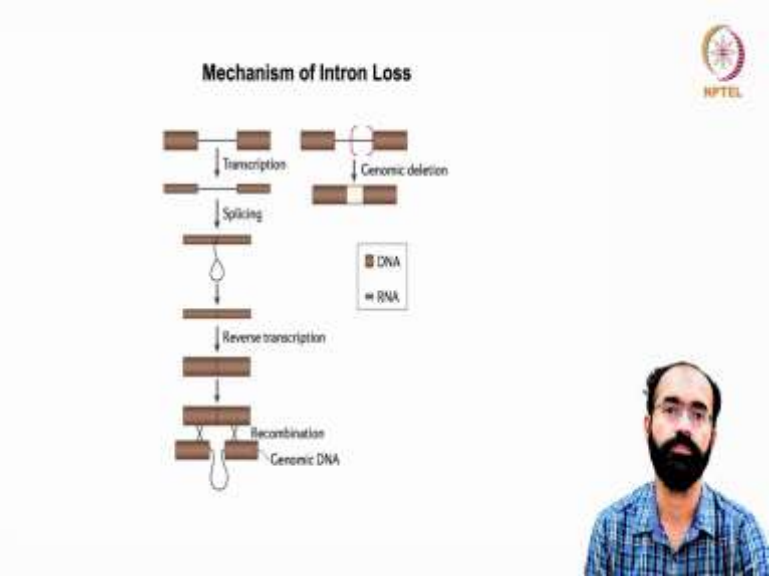
(Refer Slide Time: 24:07)



But the question is, how do we gain or how do we lose intron? So, mechanism of intron gain and loss is not fully understood. Because nobody can see it in real time we can only speculate.

So, we can assume that recombination within intron facilitated full length genome this is a feature. We know that the more recombination that can more recombination that is taking place between genes or between DNA fragments we can say more introns are possible.

(Refer Slide Time: 24:42)



But how are we going to explain it? Let us see mechanism of intron loss. This is an exon, this is another exon, this is intron and transcription takes place this is fat because it is it representation of the DNA. When transcription take place, it have a single stranded exon and exon and then intron and this splicing takes place and you end up getting exon fusion and intron is removed.

Now, you have two exons fused together. What if undergoes reverse transcription? If a virus normally we do not produce any reverse transcriptase. If a virus infection is there. So, our genes are also vulnerable for reverse transcription. Reverse transcription takes place and it can recombine back into the same genome into the genomic DNA.

Because it is now not having the intron because its already spliced and now it got converted into DNA. And DNA can recombine into the same site where it originated and if it gets integrated then this intron because genome had the intron can be lost.

And another way of losing of intron is a genomic deletion in the intron itself. So, this is the same sequence you have exon, exon and intron. A portion of intron if it is deleted you can lose part of the intron that is partial loss of the intron or it can even cause complete loss of the intron simply by genomic deletion which can happen because introns do have lot of repeat sequences.

So, it can be lost while assembling back sometimes you can lose it. And as a result, you can lose part of the intron. We will see about the intron loss and gain more in detail in the next class.

Thank you.