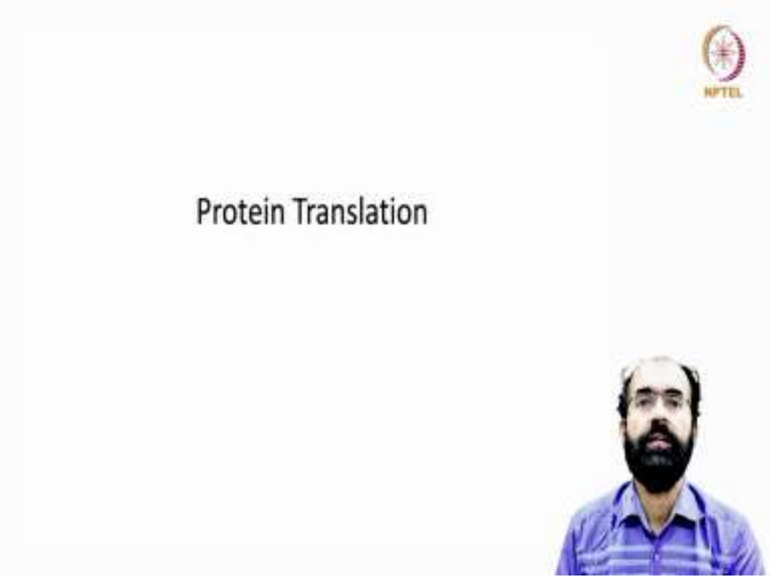
RNA Biology
Prof. Rajesh Ramachandran
Department of Biological Sciences
Indian Institute of Science Education and Research, Mohali

Lecture - 61
Epitranscriptome and Protein Synthesis: Roles of RNAs in Translation

(Refer Slide Time: 00:22)


Protein Translation

Hello everyone, welcome back to another topic of RNA Biology. So, we were learning about the importance of RNA modifications that is essential for the survival of the RNA, stability of the RNA and pushing this RNA into translation function. And many a times that is essential and many a times it can overdo than what is required and can lead to complications. So, now let us understand the importance of and the mechanism of protein translation itself.

We all know the work force in the cell are proteins. No cell can survive without active participation of several proteins. So, on an average around 7000 genes contribute for the housekeeping functions of a given cell, housekeeping functions means those functions that is performed in every cell.

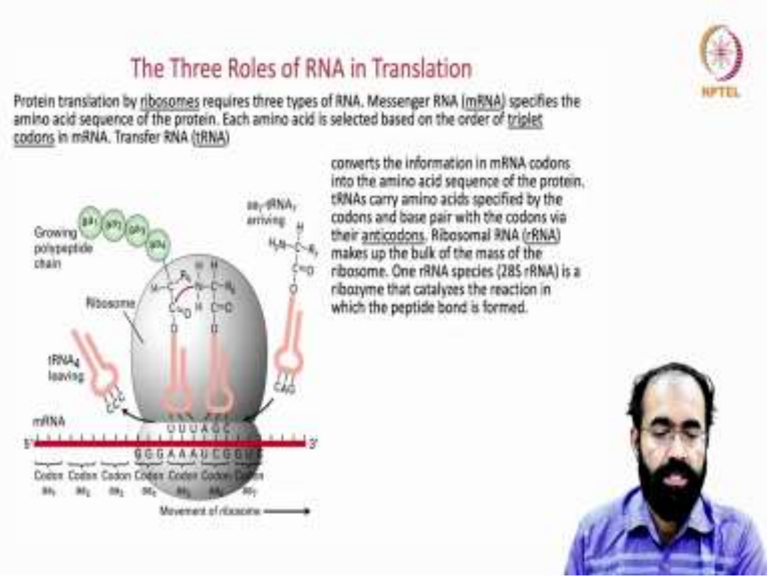
And a bunch of few thousands of genes will do the tissue specific function means liver will do liver specific function, kidney will do kidney specific function, brain will do brain specific function; spleen will do spleen specific functions.

And this is regulated via the tissue specific transcription factors. And we usually classify the transcription factors into two broad category, one is general transcription factor and the other category is the specific or tissue specific transcription factors. So, general transcription factors are mainly taking care of the transcription function. They are the work force; they are the workers that will produce RNA from the DNA.

Whereas specific transcription factors are those which will guide the system, the guide the cell to where to start the transcription, which all genes to be transcribed in this given tissue and how long this should continue. Should I, should it continue forever or should it be done only for a specific window? So, this is what normally done by specific transcription factors.

No matter general or specific, these are all proteins. So, proteins act as transcription factors, enzymes, structural proteins, etcetera. So, even to produce RNA, you need proteins and these RNA, many mRNA's will give rise to proteins. So, let us understand the mechanism of protein translation.

(Refer Slide Time: 03:12)



The Three Roles of RNA in Translation

Protein translation by ribosomes requires three types of RNA. Messenger RNA (mRNA) specifies the amino acid sequence of the protein. Each amino acid is selected based on the order of triplet codons in mRNA. Transfer RNA (tRNA) converts the information in mRNA codons into the amino acid sequence of the protein. tRNAs carry amino acids specified by the codons and base pair with the codons via their anticodons. Ribosomal RNA (rRNA) makes up the bulk of the mass of the ribosome. One rRNA species (28S rRNA) is a ribozyme that catalyzes the reaction in which the peptide bond is formed.

So, there are three roles of RNA in the translation. We have kind of discussed that the three roles in the previous classes ever since this course began. So, the protein translation by ribosomes require three types of RNA. First one is the mRNA or the messenger RNA, which contains the message from the genome for a specific gene into the cytoplasm.

We call it as mRNA and this specifies the amino acid sequence of the protein that is going to be made from this RNA. So, each amino acid is selected based on the order of triplet codons in the mRNA, like each amino acid require for the each incorporation of the amino acid require three bases and we call it as a codon rather triplet codon.

And then the transfer RNA, transfer RNA brings in individual amino acids. So, total 64 codons are possible in a four base system, like we have four bases in an RNA, A, U, G and C. So, what the process of translation means, it converts the formation of in the mRNA to codons into amino acid sequence of the protein. Say if you have 300 codons in a mRNA, it will give rise to 300 amino acid, 300 codon means 900 bases, 300 into 3.

So, you end up getting a 900 base and talking about the start codon to stop codon and this 900 base can potentially add 300 amino acids. So, that is the concept into amino acid sequence of the protein. So, tRNAs carry amino acid specified by the codons and base pair with the codons via their anticodon. So, anticodon is another three base, which is complementary to the codon. So, anticodon is present on the tRNA. The codons are present on the mRNA.

So, the anticodons and codons pair and that will deliver one amino acid. So, just like if you send a letter with an address, the postman will do the matching, he will read the address on the letter and he also read the address what is of your house and then deliver the letter. This is somewhat similar how a anticodon and codon pair and the amino acid delivery in the site of protein synthesis.

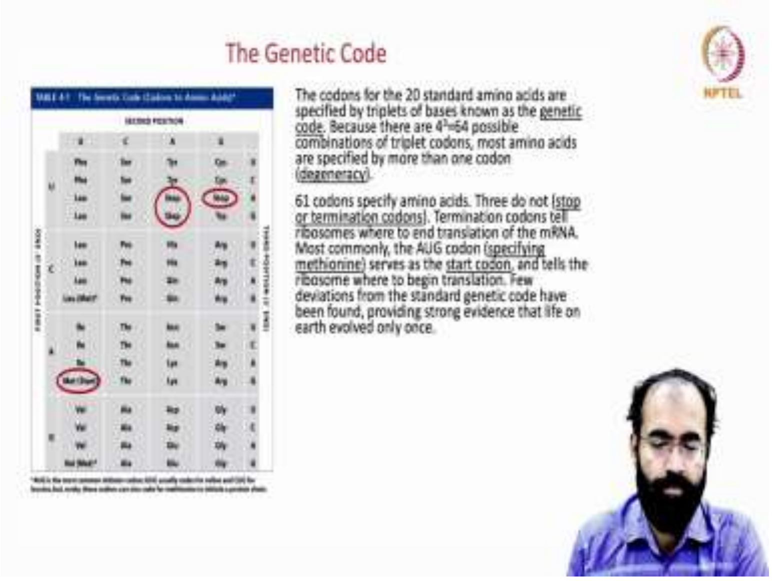
So, the ribosomal RNA that is rRNA that will constitute the bulk of the ribosome. Although ribosome is a RNA plus protein complex, the ribosomal RNA does lot of the catalytic role. So, one RNA species that is the 20S, 20S ribosomal RNA is a ribozyme. We have learnt a lot about that, ribozyme that catalyzes the reaction in which the peptide bond is formed.

So, remember the 28s ribosome, which is a ribozyme is the one which is contributing to the formation of the. So, called peptide bond, peptide bond is the bond between two amino acid. Amino acid A, amino acid B, amino acid C, they are paired together between A and B, there is a peptide bond between B and C, there is another peptide bond. So, like that it continues. So, you can see here this red color one is the mRNA and it has got codon 1, codon 2, codon 3, codon 4, codon 5, codon 6, codon 7.

Each one corresponds to amino acid 1, amino acid 2, amino acid 3, amino acid 4 and you can see here the each of the tRNA when it is done using because it enter through the you know A site and the peptide bond is formed in the P site and it is the eliminated through the exit site E site. So, this is a removed a tRNA, this is the ribosome two bubbles and you can see the middle one where the peptide bond is actually formed that is the one which is having the growing amino acid chain.

So, in this codon 1 initially this ribosomes were here and it had the first amino acid then second, then third, then fourth. So, corresponding to the fifth one this tRNA has brought in the fifth one and at here corresponding to the fifth codon you have this tRNA and the sixth one and seventh one is arriving. So, this will continue until the total number of codons are completed. So, that is the basic principle of the protein translation.

(Refer Slide Time: 08:09)



The Genetic Code

The codons for the 20 standard amino acids are specified by triplets of bases known as the genetic code. Because there are 4^3=64 possible combinations of triplet codons, most amino acids are specified by more than one codon (degeneracy).

61 codons specify amino acids. Three do not (stop or termination codons). Termination codons tell ribosomes where to end translation of the mRNA. Most commonly, the AUG codon (specifying methionine) serves as the start codon, and tells the ribosome where to begin translation. Few deviations from the standard genetic code have been found, providing strong evidence that life on earth evolved only once.

So, let us see the genetic code. We have seen this picture earlier also. So, total possible codons are 64. The codons for 20 standard amino acids are specified by triplets of bases known as the genetic code. And we have discussed about this earlier also and Har Gobind Khorana is one of the Indian origin scientist, who was contributing contributed significantly to the discovery of genetic code and he was awarded Nobel Prize.

So, because there are 4 bases and you group them into 3, 4 raise to 3 is equal to 64 basically mean 4 into 4 into 4 possibilities are there. So, total 64 possible combinations of triplet codons are possible and most amino acids are specified by more than one codon and we refer to this phenomenon as degeneracy or we call it as degenerate codon.

So, what does it mean? One amino acid, because there are only 20 amino acids, we do not have 64 amino acids to have identity. So, what it happens? One amino acids, can have starting from 1 to 6 some amino acids can have say 1 codon, some have 2 codon, some have 4 codon, some have 6 codon.

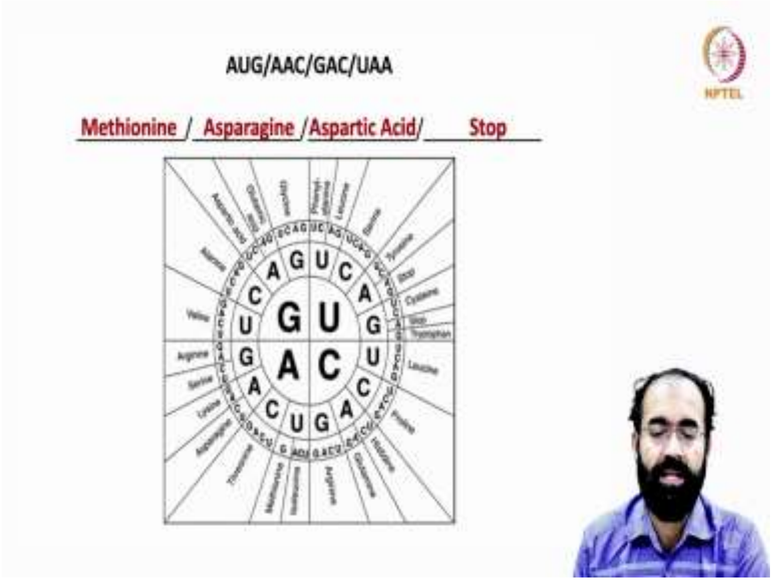
So, this is the range based on which amino acids you are talking about only 61 codons are meant for bringing in amino acids whereas, 3 of them act as the stop codon, UAA, UAG, UGA. So, these are the stop codons. So, 3 do not bring in any amino acids and we call them as termination codons or stop codons.

So, termination codons tell the ribosome where to end the translation of the mRNA. Most commonly AUG codon is the starting codon and it specifies methionine and there is no second codon available for methionine. So, methionine has got only one codon and that is AUG and that is the start codon for every protein; serves as the start codon that is AUG serves as the start codon and tells the ribosome where to begin the translation process.

Few deviations from the standard genetic code have been found providing strong evidence that life on earth evolved only once. Because a lot of organisms starting from virus to human, virus, bacteria, fungus, human everywhere the codons are constant and same codon brings in same amino acid. You can never find a given codon bringing in another amino acid.

So, this is one of the strongest support to believe that the life originated only once. Because if it was originated multiple times, it is impossible to have the same codon present in all the attempts. Because there is no reason, there is no rule to be followed that this amino acid has to be brought in with only with this codon. At least you should have had a variation in the codon and even you do not see that. So, this is another way of presenting.

(Refer Slide Time: 11:38)



AUG/AAC/GAC/UAA

Methionine / Asparagine /Aspartic Acid/ Stop

So, one table is here you have got the first position base, second position base and the third position base. That means first position U and the second position is U and the third is U. So, here you will say U, U, U like that. So, these are the short forms of amino acid Phe stands for Phenylalanine, Leu stands for leucine and Ile stands for Isoleucine, Thr stands for Threonine like that.

So, the names of amino acids are not important at this moment, but understand 20 different types of amino acids are there. Another way of presenting the genetic code is this way. This is the first one, second one and the third one. So, G, C, G one codon, G, C, A another codon, G, C, C another codon, G, C, U another codon, they all bring in alanine and like that you can go valine, arginine, serine like that.

So, for arginine you have A, G, G, A, G, A arginine, but arginine can have elsewhere also. See here also you have C, G, G, arginine, C, G, A arginine, C, G, C arginine, C, G, U arginine. So, arginine have got 4 here, 2 here. So, total 6 arginine has. So, only methionine and tryptophan, 2 amino acids have got only 1 codon. Rest all of them have at least 2 codons are there. So, we should also understand in bacteria, plants and mammals or the animals. Plants, bacteria plants prefer a certain set of codons, but they follow the same rule.

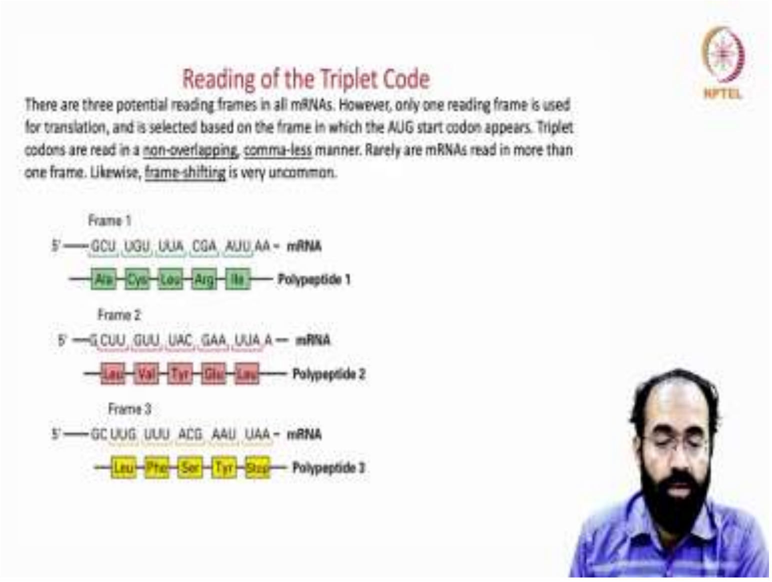
Certain set means those codons ending with A or U is preferred whereas, in animals the codons ending with, ending means the third base codon what I mean is A, B, C, 3 bases are there, ending with G or C is preferred. This does not mean that plants do not have a codon ending with G or C is missing or this also does not mean that animals do not have codons ending with A or U. Both are there we are talking about preference.

Like say for example, if arginine have got a codon A, G, G this is preferred compared to A, G, A because it is ending with A. Something like that you can see C, G, G is preferred, C, G, A is less preferred or C, G, C is preferred, C, G, U is less preferred in animals.

So, plants it is the other way round or in bacteria, but all codons are present all codons are present in all animals. There is no codon is missing like of the 61 codon entire 61 codons are represented in all living forms. There is no single living form is available which is lacking any one of this codon which is not served by a tRNA. So, that is not correct.

So, you can see methionine, asparagines, aspartic acid and stop codon. So, this is an example like you start with methionine and you can have any amino acid do not think that it has to asparagine or aspartic acid and eventually it will come to a stop codon. This is the easiest way of depicting a translation corresponding to these codons. AUG brings methionine, AAC brings asparagine, GAC brings aspartic acid, UAA brings as stop codon. But keep in mind no mRNA will have just 3 codons or 3 amino acids. So, that is not a mRNA. So, it will have much more number.

(Refer Slide Time: 15:27)



Reading of the Triplet Code

There are three potential reading frames in all mRNAs. However, only one reading frame is used for translation, and is selected based on the frame in which the AUG start codon appears. Triplet codons are read in a non-overlapping, comma-less manner. Rarely are mRNAs read in more than one frame. Likewise, frame-shifting is very uncommon.

So, now reading of the triplet cod. So, there are 3 potential reading frames in all RNA. And because of this because no one says ok, this is the first base and this is the second base. No there is no way RNA have a first base and RNA have got a last base, but not for the open reading frame or where is the start, start also we know AUG. But there will be several AUG who will tell you that which AUG you should start?

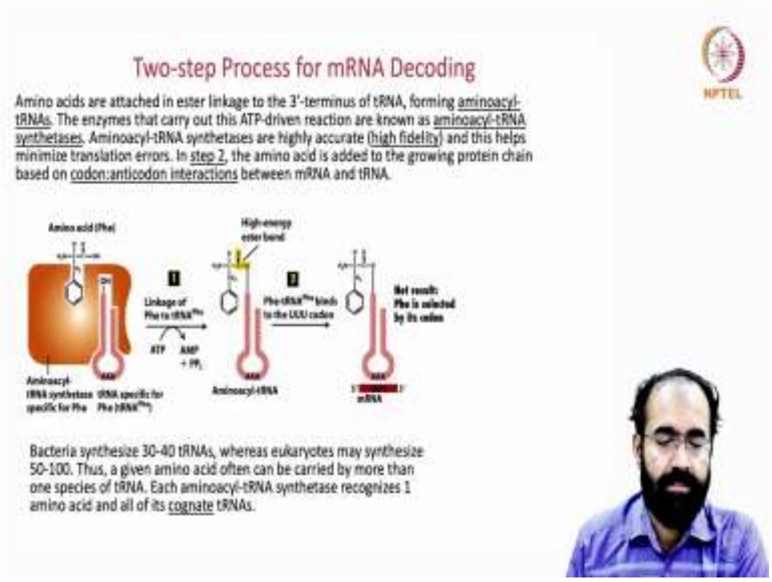
So, this is always possible and 3 potential reading frames in all RNA's. However only one reading frame is used for translation and it is selected based on the frame in which the AUG codon appears. And you should also understand in the same RNA there may be 3 or 4 AUG's will be there. So, not that every AUG it will start translation even if it starts it will immediately find a stop codon, say maximum it will go to around 6 or 7 amino acids. It will find a stop codon, hence that is not preferred.

So, the right AUG when it is starting then it will find a bunch of amino acids and it will reach until the stop codon and the entire protein is translated. So, the system make sure that any out of frame reading happens it will not be wasted wasting the resources in the cell by producing unnecessary protein. So, it will immediately marked for degradation because it simply makes around 7 or 8 amino acids only.

So, the triplet codons are read in a non overlapping or in a comma less manner. So, rarely there are mRNA's that read in more than one frame. Likewise, the frame shifting is very uncommon. Say for example, like you can see here frame 1, frame 2 and frame 3 are mentioned here. So, let us see GCU is one frame and UGU another codon. And so, this is one frame starting from GCU. So, now this GCU is G is not counted whereas, CUU is counted that means, this CU and U. So, there is a frame shift and it continues like that.

CUU, GUU like that it is continuing. And what is the other frame? It is skipping two base like this is GCU, G and C is skipped here. So, UUG that is UUG is taken. So, UUG, UUU, ACG, AAU, UAA like that it is continuing. So, we can say each frame has the potential to give a polypeptide, polypeptide 1, polypeptide 2 polypeptide 3, but in reality, the right one is only going to be completing a substantial stretch of amino acid. The other frames often end up in a quick stop codon or a very frequent stop codon.

(Refer Slide Time: 18:26)



Two-step Process for mRNA Decoding

Amino acids are attached in ester linkage to the 3'-terminus of tRNA, forming aminoacyl-tRNAs. The enzymes that carry out this ATP-driven reaction are known as aminoacyl-tRNA synthetases. Aminoacyl-tRNA synthetases are highly accurate (high fidelity) and this helps minimize translation errors. In step 2, the amino acid is added to the growing protein chain based on codon:anticodon interactions between mRNA and tRNA.

Bacteria synthesize 30-40 tRNAs, whereas eukaryotes may synthesize 50-100. Thus, a given amino acid often can be carried by more than one species of tRNA. Each aminoacyl-tRNA synthetase recognizes 1 amino acid and all of its cognate tRNAs.

So, let us understand more about the two step process of mRNA decoding means protein translation is nothing but mRNA decoding. So, amino acids are attached in ester linkage to the 3 prime terminus of tRNA forming amino acyl tRNAs. The enzymes that carry out this ATP driven reaction are known as amino acid tRNA synthetase. So, tRNA produced in the nucleus that undergoes the maturation in the cytoplasm and it also gets amino acylated. So, that this tRNA gets ready.

So, amino acyl tRNAs synthetases are highly accurate because it cannot randomly catch hold of some amino acid and add on to a given tRNA. Because this amino acid must correspond with its anticodon because each tRNA should have the correct anticodon sequence so that this amino acid will be delivered in the right codon of the mRNA during protein synthesis.

And this is called high fidelity means there is no scope for error. If a tryptophan amino acid should be delivered to the tRNA that contains the anticodon sequence for the tryptophans codon. Otherwise the tryptophan will end up elsewhere. So, you do not want that to happen, that is why the high fidelity of the amino acyl tRNAs synthetase are very important.

And this helps to minimize translation errors. In step 2 that happens in step 1, step 2 the amino acids is added to the growing protein chain based on codon anticodon interactions that takes place in presence of the ribosome subunit between the mRNA and the tRNA.

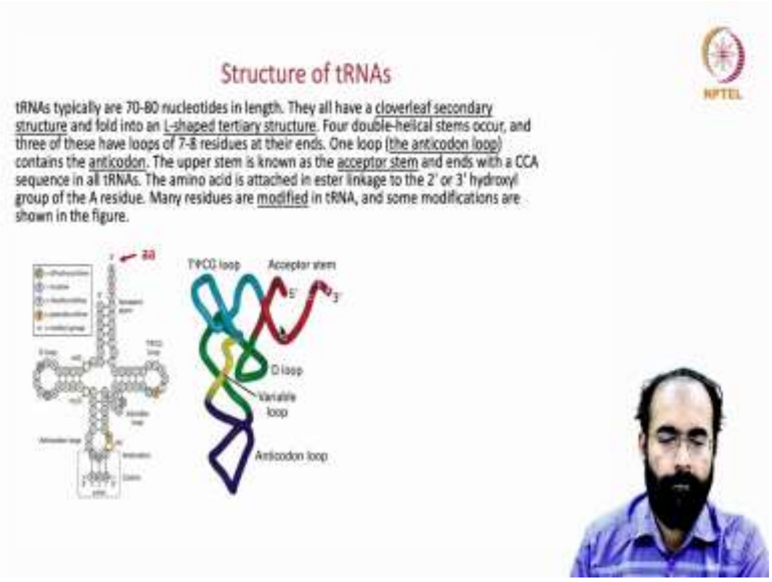
That is what you can see here. Amino acyl tRNA synthetase catch hold of an amino acid and catch hold of the tRNA and it creates the linkage of the tRNA with the amino acid and this tRNA has to pair in the right place.

So, this high energy ester bond is broken only inside a ribosome where the protein synthesis is taking place or the peptide bond is taking place. So, the aminoacyl tRNA synthetase make sure the charging or the loading of the tRNA with the amino acid. Bacteria synthesize around 30 to 40 tRNAs whereas, eukaryotes synthesize around 50 to 100. Thus a given amino acid often can be carried by more than one species of tRNA.

Each amino acyl tRNAs synthetase recognize one amino acid and all of its cognate tRNA. Say for example, let us talk about the amino acyl tRNA synthetase of arginine. We know arginine has got 6 codons. So, naturally it will have 6 tRNAs with the corresponding anticodons.

So, the arginine amino acyl tRNAs synthetase can recognize all the 6 tRNAs that contains the corresponding anticodon. So, one amino acid tRNAs synthetase can happily deliver the arginine amino acid to all the 6 different types of tRNAs corresponding to the codons of these amino acid.

(Refer Slide Time: 22:20)



Structure of tRNAs

tRNAs typically are 70-80 nucleotides in length. They all have a cloverleaf secondary structure and fold into an L-shaped tertiary structure. Four double-helical stems occur, and three of these have loops of 7-8 residues at their ends. One loop (the anticodon loop) contains the anticodon. The upper stem is known as the acceptor stem and ends with a CCA sequence in all tRNAs. The amino acid is attached in ester linkage to the 2' or 3' hydroxyl group of the A residue. Many residues are modified in tRNA, and some modifications are shown in the figure.

So, now let us look into the structure of tRNAs. tRNAs are roughly 60, 70 to 80 nucleotides in length that is there. It is much longer to start with which undergoes maturation, splicing etcetera. And they all have a clover leaf secondary structure and fold into an L-shaped tertiary structure. So, secondary structure is clover leaf and tertiary structure is L-shaped. And four double helical stems occur and three of these have loops of 7 to 8 residues at their ends.

And some of the CCA modification we discussed in the earlier classes which is a post-transcriptional modification. Every amino acid should have a CCA end in the 3 prime end. And one loop that we call it as anticodon loop that is this area. Anticodon loop pairs with the codon. And this contains the anticodon sequence a 3 base sequence corresponding to the complementary sequence of the codon.

So, the upper stem is known as the acceptor stem and it ends with a CCA sequence in all tRNAs. Why it is same? Because all tRNAs come from different genes and none of them have the CCA to start with because it is done by a post-transcriptional modifications. So, all tRNAs have a 3 prime end CCA. So, the amino acid is attached in ester linkage to the 2 prime or 3 prime hydroxyl group of an A residue.

A residue means adenine residue. Many residues are modified in tRNA and some modifications are shown as you can see here in below. So, this modification some of them we have already discussed this called RNA editing that allows with minimum you can perform maximum.

That means some modifications are changing into inosine. An inosine is a neutral base. It does not, it pair equally well with all the 4 bases A, G, C or U it neither attracts nor repels. So, in this way some animals will be able to cater more codons with less number of tRNA diversity or tRNA diversity is low you can edit this anticodon arm and you can cater those codons also.

And some of these modifications are providing like pseudo-uridine etcetera. And psi loop these modifications protect the tRNA from some nucleases etcetera. So, we should understand that the tRNA will have a dedicated region that is the anticodon arm and a dedicated region for the binding of the amino acid. So, this binding ensures that the amino acids are strongly attached and they are delivered in the right place and right time to the required targets.

So, we should appreciate the beauty of the protein translation with all its merits going to various enzymes. They are specific like amino acyl tRNA synthetase that are very specific and also various enzymes that creates the RNA modifications of certain bases and lot more enzymes come into picture to provide specificity to the protein synthesis. So, we will learn more about the protein synthesis in the next class.

Thank you.