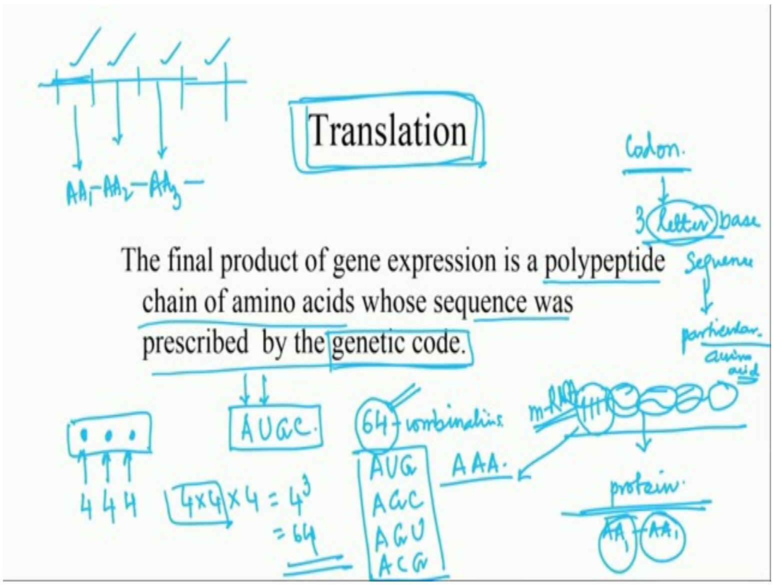
Organic Chemistry In Biology And Drug Development
Prof. Amit Basak
Department of Chemistry
Indian Institute of Technology, Kharagpur

Lecture - 30
Central Dogma: DNA Replication, Transcription and Translation (Contd.)

Welcome back to this course. In the last session, we have discussed about the process of formation of the RNA from a double stranded DNA. This process was called transcription which is a part of the central dogma biology. The next step is synthesis of protein from RNA and the process is called translation.

(Refer Slide Time: 01:03)



Translation

The final product of gene expression is a polypeptide chain of amino acids whose sequence was prescribed by the genetic code.

In translation process, the final product of gene expression is making a polypeptide chain of amino acids whose sequence was prescribed by the genetic code. So, there are now certain things which you need to clarify. One thing is the genetic code, another is m-RNA which is actually into the protein.

Amino acids are linked together to one after another in a protein. So, which amino acid will be there? That will depend on the sequence of bases in the m-RNA, and that is what is called the genetic code. So, genetic code is basically a 3 letter code. We have this 4 bases AUGC. Out of this 4 bases, you can have 64 combination by permutation and combination.

For example, you can have AUG/AGC/ACG and you can have these 64 combinations.

For the first position, you have 4 bases. So, you have four possibilities to fill up the first position. Also in the second position, you have four possibilities because you can have repetitions. Similarly for the third position you can have four possibilities. So, the total number of possibilities is $-4*4*4 = 4^3 = 64$.

So, these three letters actually code for a particular amino acid. There are 20 amino acids present in the nature.

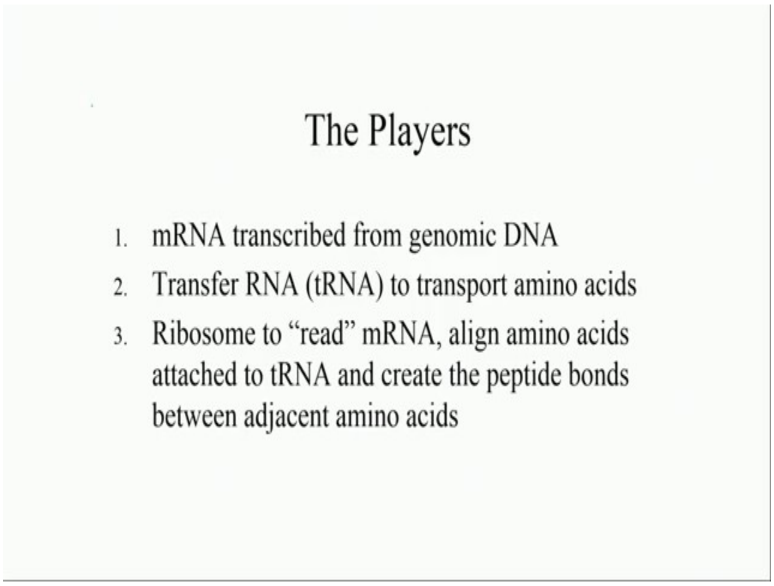
We need at least 20 genetic codes to translate amino acids. If you take a two letter code your number of genetic codes will be 16. That does not encompass the whole 20 amino acids. So, nature had no other option to bring in a third letter to cover the entire amino acid spectrum.

These 3 letters codes are called codon. You have many extra codes because you have 64 codons. A particular amino acid is represented by particular codon.

Out of these 64 codons, there are 3 codons which are called stop codons or nonsense codons. Nonsense means that does not make any sense i.e. that does not code for any amino acid. So, these are called stop codons. And the rest 61 are coding for different amino acids.

In translation process, protein is synthesized from mRNA. The primary structure of protein represents different amino acids. Different proteins have different amino acid sequence. The amino acid sequence is dictated by the sequence of codons in the m-RNA. So, you can divide m-RNA into several codons. If it is translated into protein, you will get amino acid sequence depending upon the codons present. So, this way the amino acid sequence is varied by varying the different codons s in the m-RNA. Remember the codon is present in the m-RNA.

(Refer Slide Time: 09:13)



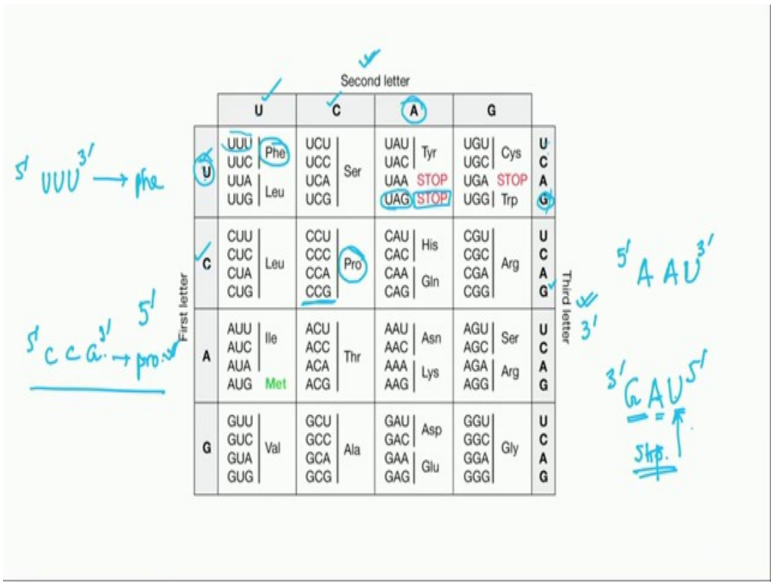
The Players

1. mRNA transcribed from genomic DNA
2. Transfer RNA (tRNA) to transport amino acids
3. Ribosome to "read" mRNA, align amino acids attached to tRNA and create the peptide bonds between adjacent amino acids

64 codons are represented in a table by a systematic way like a periodic table. The table containing the all the 64 codons and reading that you can tell which amino acid is coded by which codon. Because it is impossible to remember the codons that represents a particular amino acid.

Rather than memorizing this it is better to check the codon that codes for this particular amino acid.

(Refer Slide Time: 10:43)



	Second letter				
First letter	U	C	A	G	Third letter
U	UUU Phe UUC UUA Leu UUG	UCU Ser UCC UCA UCG	UAU Tyr UAC UAA STOP UAG STOP	UGU Cys UGC UGA STOP UGG Trp	U C A G
C	CUU Leu CUC CUA CUG	CCU Pro CCC CCA CCG	CAU His CAC CAA Gln CAG	CGU Arg CGC CGA CGG	U C A G
A	AUU Ile AUC AUA AUG Met	ACU Thr ACC ACA ACG	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G
G	GUU Val GUC GUA GUG	GCU Ala GCC GCA GCG	GAU Asp GAC GAA Glu GAG	GGU Gly GGC GGA GGG	U C A G

5' UUU 3' → phe

5' CCG 3' → pro 5'

5' AAU 3'

3'

3' GAU 5'

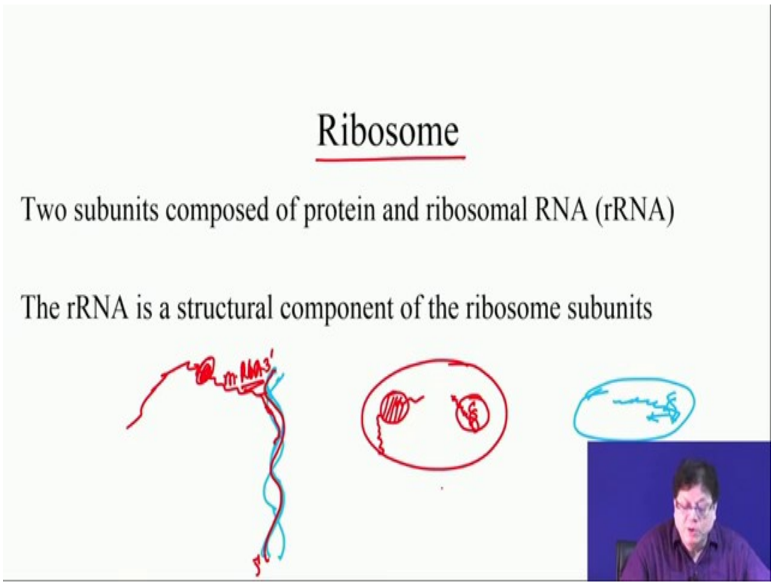
Stp.

How this table is read? See you see all this UCAG on this right side, then UCAG on the top side, and then you have several the others. You have to always specify which the 5 prime end is and which the 3 prime end is. Now, the first letter is starting from 5 prime end and the third letter is ending at 3 prime end. For the code 5’-CCG-3’, first letter is C, second letter is C and third letter is G.

For the code 3’-GAU-5’, first letter is U, second letter is A and third letter is G. So, it is actually 5’-UAG-3’.

UAG is a stop codon. As soon as you reach that UAG the protein synthesis mechanism will stop.

(Refer Slide Time: 14:39)



Ribosome

Two subunits composed of protein and ribosomal RNA (rRNA)

The rRNA is a structural component of the ribosome subunits

In prokaryotes, you do not have any compartments. Suppose this is your DNA. It will be transcribed into the m-RNA. The m-RNA is coming out into the cytosol. So, the protein synthesis can take place before the whole m-RNA comes out of the DNA machinery.

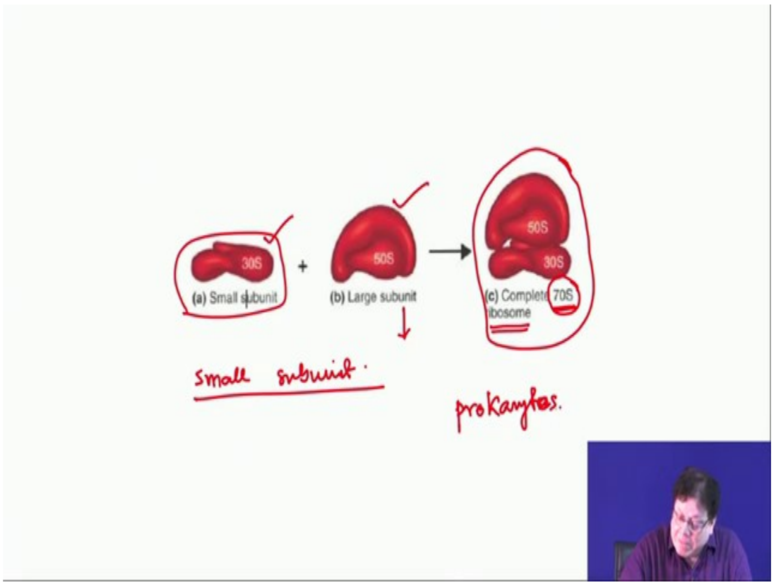
Then the synthesis is not complete yet because there is no separate compartment.

There is this protein making factory which is called ribosome and it is also present in the cytosol. So, the ribosome will start making the proteins. That means, before the m-RNA synthesis is complete in prokaryotes, the ribosome can start working to synthesize the protein from m-RNA. the proteins, because this is the m-RNA.

But in eukaryotes that does not happen because all eukaryotes have a separate nucleus where DNA is present.

So, when the RNA is transcribed it will more or less remain here. Once it is complete then it is released and goes into the ribosome. This is suppose the ribosome and the protein is now synthesized depending on the codon sequence of the m-RNA. It cannot go to the ribosome in straight way because in eukaryotes whatever m-RNA is prepared that needs to be processed. Let us first talk about the prokaryotic system.

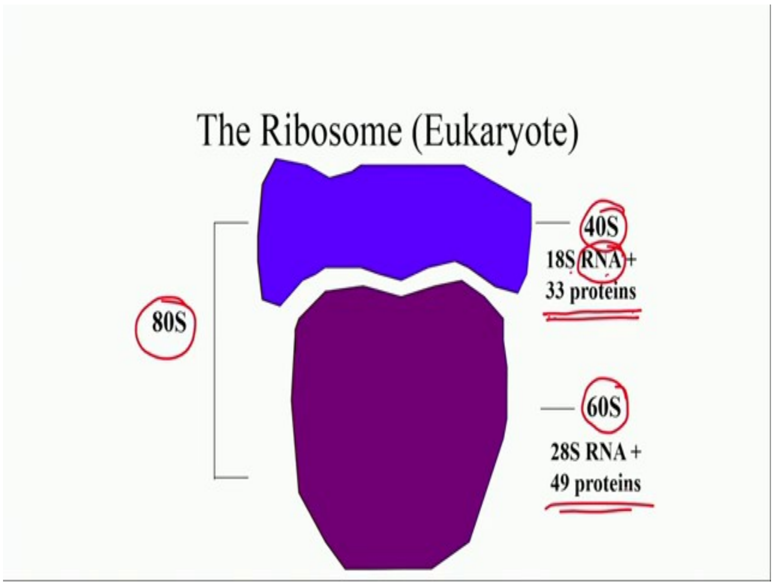
(Refer Slide Time: 18:39)



Ribosome actually has two sub units. One is called a small subunit whose which is the 30 S subunit. S is the Svedberg unit. It depends on the sedimentation rate of the system or the molecule.

So, this is a 30 S subunit and there is a 50 S subunit. There is a small subunit and a big subunit. This is the large subunit. The protein synthesis is taking place when both are actually held together and that is called the complete ribosome where the protein synthesis ultimately takes place. These two things are held and the overall weight will not be 50 plus 30. It is not just addition. If you study on sedimentation you will see that that becomes 70S.

(Refer Slide Time: 20:35)

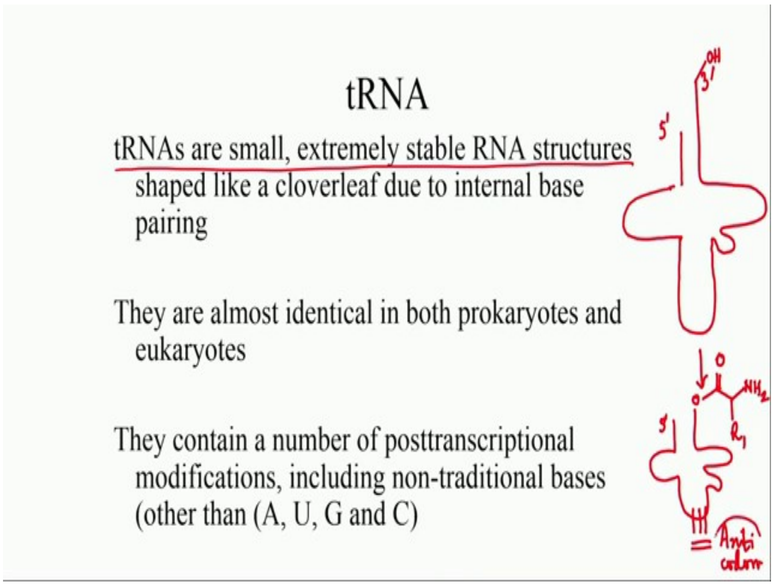


The Ribosome (Eukaryote)

Ribosome of prokaryotes is little bit smaller than the eukaryotes. Eukaryotes ribosome is 40S plus 60S. It has also have a small and a large subunit. The protein synthesis is done when the two are held together. So, the total ultimately becomes 80 S. So, the eukaryotes have 80 S ribosome consisting of 40 S small unit and a 60 S small unit. In the eukaryotes you have 70 S ribosome, and that has got a 30 S and a 50 S subunit. Ribosome is made up of ribosomal RNAs and proteins. If you take this eukaryote ribosome the small subunit has 18 S RNA. This is called ribosomal RNA and it contains 33 proteins. The large one has 28 S RNA and 49 proteins.

All these are held together and form the ribosome complex. Ribosome is nothing but combination of proteins and ribosomal RNA. This is one of the most important machinery in the cell because you are making the proteins which we require to sustain the living system. The DNA has directed via RNA to synthesize the proteins. The proteins do the entire job that is necessary.

(Refer Slide Time: 22:41)



tRNA

tRNAs are small, extremely stable RNA structures shaped like a cloverleaf due to internal base pairing

They are almost identical in both prokaryotes and eukaryotes

They contain a number of posttranscriptional modifications, including non-traditional bases (other than (A, U, G and C)

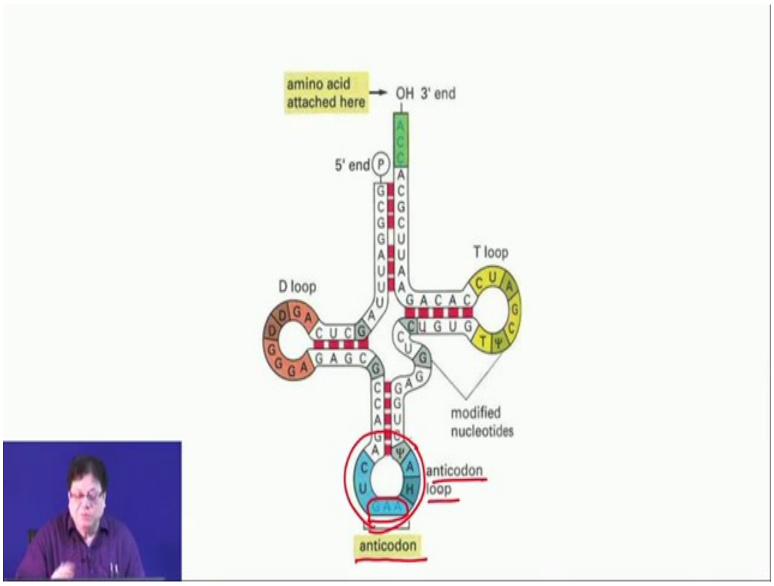
You know the template strand sequence and from that you can write the sequence of the messenger RNA. So, you can tell the codon sequence in the m-RNA. You can write the primary sequence of the protein. The ribosomal RNA plays an important role and is present in the ribosome which is the protein making factory. The duty of the tRNA is to bring the amino acids one after another to grow the protein chain.

How does the tRNA look like? tRNA has a structure like this. These molecules may be complementary inward like the hairpin. It is a single molecule, but those folds because one portion is complementary to the other portion. The tRNA also has many complementarity. It ultimately takes a shape like this. It is a cloverleaf type of shape. Now, this is the 5 prime region and this is the 3 prime region.

In the 3 prime region there is a OH and it will be connected to amino acid through acylation. So, the amino acids are basically forming an ester bond and it is attached to the tRNA. These amino acids are actually being delivered to the growing polypeptide change by the tRNA.

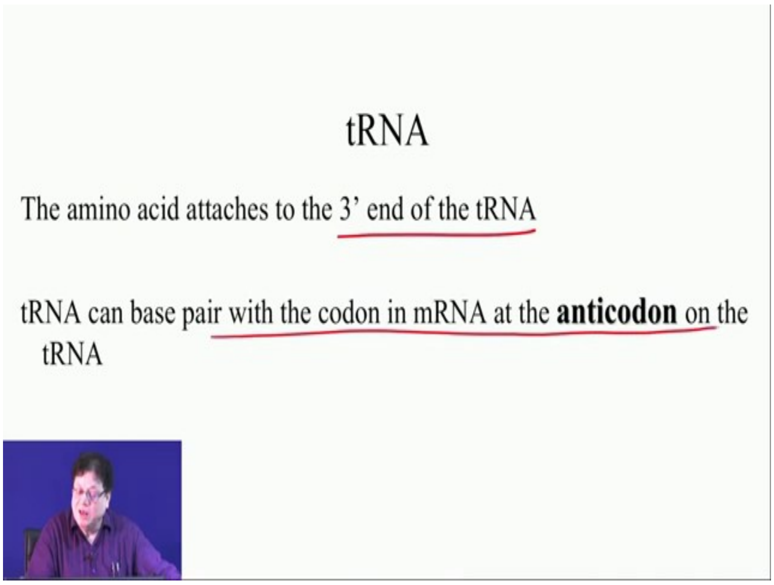
Now, tRNA has another very important aspect and that is the anticodon site. tRNA has to bind the ribosome to supply this amino acid. The position where tRNA will bind that will depend upon anti codon sequence. So, the tRNA has this acylated amino acid and it has got an anticodon basically the opposite of the codon. You can say it is the complementary of the genetic codon.

(Refer Slide Time: 26:35)



t-RNA has intra molecular hydrogen bonding like this. The 5 prime end is protected as a phosphate and the amino acids are attached at 3 prime OH. This one like GAA called the anticodon and that is called an anticodon loop.

(Refer Slide Time: 27:17)



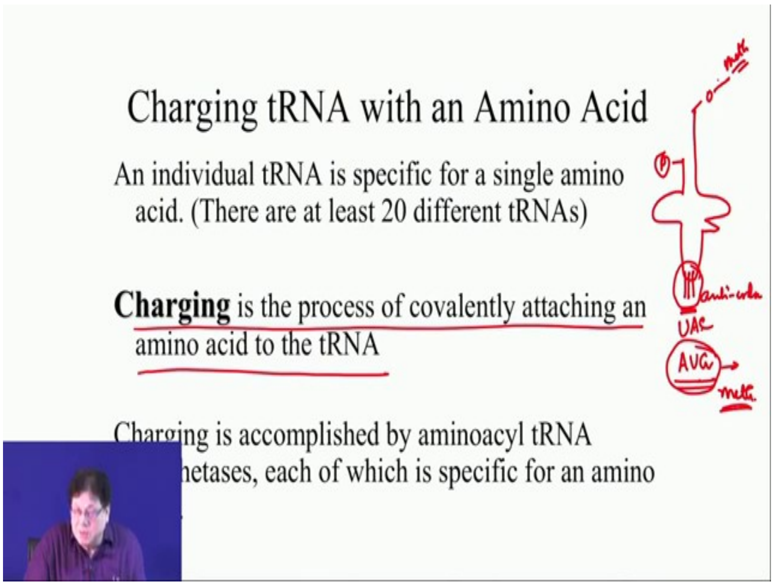
tRNA

The amino acid attaches to the 3' end of the tRNA

tRNA can base pair with the codon in mRNA at the **anticodon** on the tRNA

So, now let us go to the translation machinery. tRNA can do base pair with the codon in m-RNA at the anticodon on the tRNA. Let us see, what does it mean?

(Refer Slide Time: 27:33)



Charging tRNA with an Amino Acid

An individual tRNA is specific for a single amino acid. (There are at least 20 different tRNAs)

Charging is the process of covalently attaching an amino acid to the tRNA

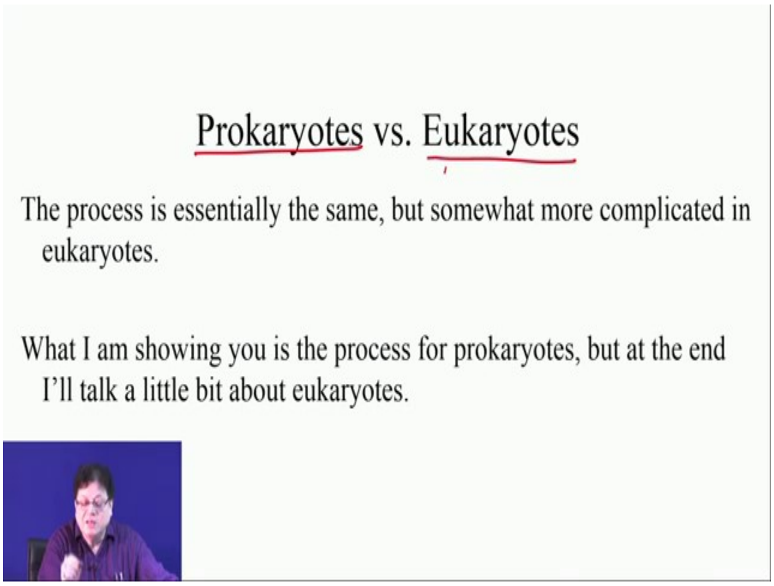
Charging is accomplished by aminoacyl tRNA
hetases, each of which is specific for an amino

Once the m-RNA is synthesized, the tRNA was free. They have to now hold up the amino acids via this acylation. So, each tRNA is covalently attached to amino acids. Now, which amino acid will be attached that depends on the sequence of the anticodon, Which amino acid will be attached here that depends upon the anticodon sequence.

Suppose the anticodon is UAC, then the codon of this will be AUG. Now, see the genetic code table. AUG codes for methionine. The amino acid will be taken up depending on the anticodon.

So, you see the anti-codon and then write codon sequence. Then go to the genetic code table to find the amino acid that is attached.

(Refer Slide Time: 30:39)



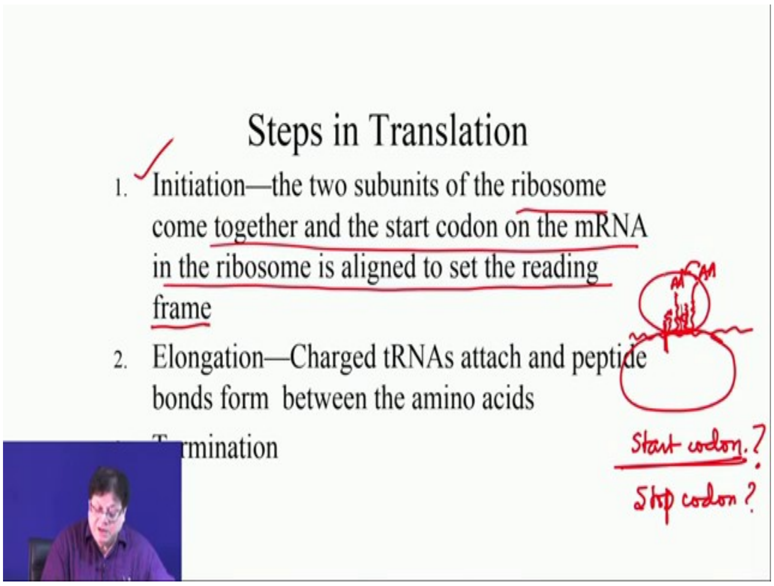
Prokaryotes vs. Eukaryotes

The process is essentially the same, but somewhat more complicated in eukaryotes.

What I am showing you is the process for prokaryotes, but at the end I'll talk a little bit about eukaryotes.

Eukaryotes are much more complicated. So, we will talk about the prokaryotes and that will be sufficient for our course.

(Refer Slide Time: 30:51)



Steps in Translation

1. Initiation—the two subunits of the ribosome come together and the start codon on the mRNA in the ribosome is aligned to set the reading frame
2. Elongation—Charged tRNAs attach and peptide bonds form between the amino acids

Termination

Start codon ?

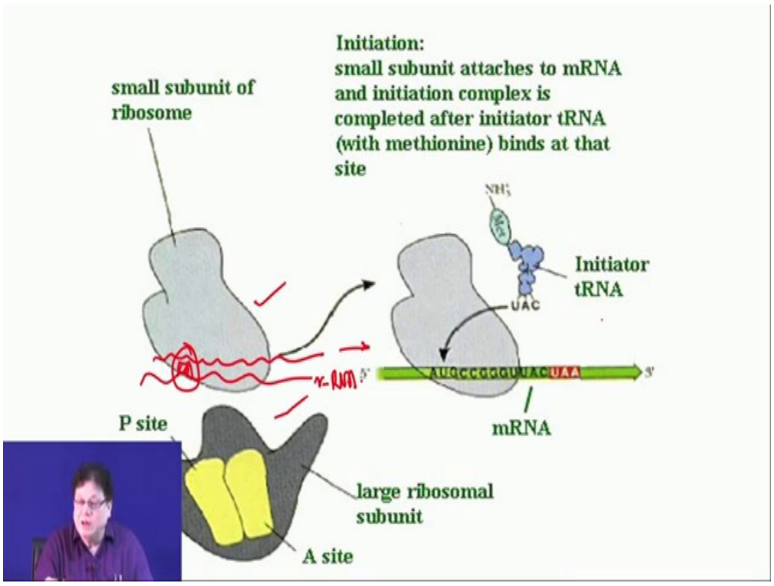
Stop codon ?

First step of translation is initiation. Two subunits of the ribosome come together and the start codon on the m-RNA in the ribosome is aligned to set the reading frame. Sequence of codons in the m-RNA is called the reading frame. You have the small subunit and the big subunit of ribosome. Your messenger RNA has been is attached to this ribosome and then the it has to start. So, there must be a start codon. Otherwise again the question comes that where does the protein synthesis start and which amino acid is added in the cell.

What is a start codon? A particular codon is start codon. So, the protein synthesis now starts. The tRNA will come here. tRNA first binds here and brings 1 amino acid here. Then another tRNA comes here and we will have another amino acid here at the 5 prime end. Now, there will be a reaction between these two amino acids.

This amino acid is transferred to that one making this tRNA free. tRNA will now leave and another tRNA will come and sit here. Then the reaction takes place like this. See the diagram.

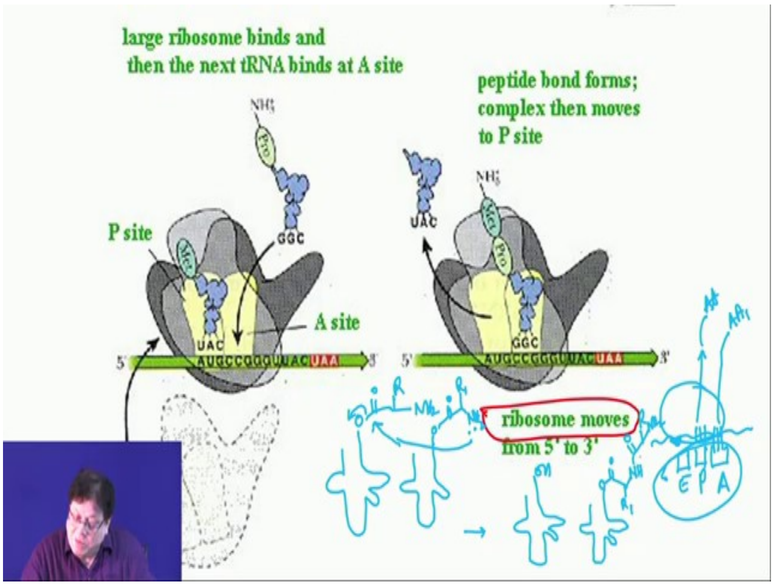
(Refer Slide Time: 33:31)



Initiation:
small subunit attaches to mRNA and initiation complex is completed after initiator tRNA (with methionine) binds at that site

Initially the m-RNA comes directly from the cytosol. There is kind of a promoter region in the ribosomal RNA that is present in this small subunit. m-RNA binds to the small subunit because there is a binding pocket in the small subunit of the ribosomal RNA.

(Refer Slide Time: 34:49)



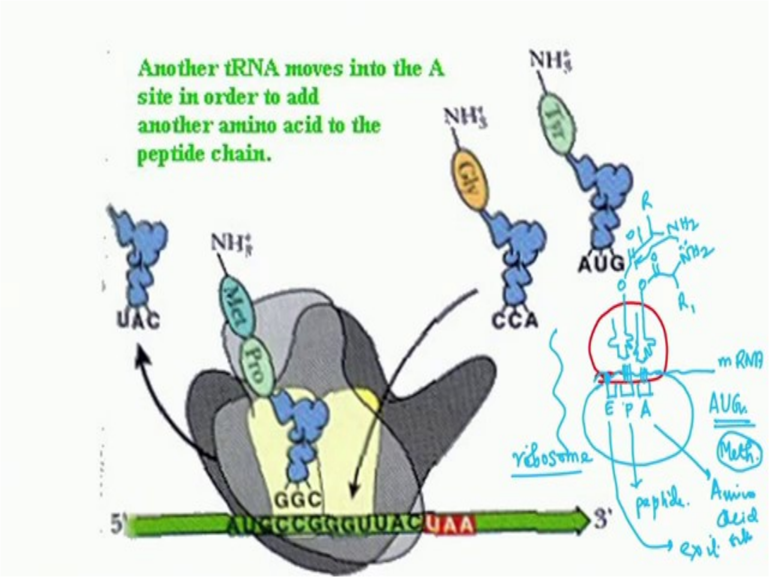
large ribosome binds and then the next tRNA binds at A site

peptide bond forms; complex then moves to P site

ribosome moves from 5' to 3'

Then the large subunit comes and forms a complex. So, this is what the complex. The complex is called the ribosome. But now you have a combination of both small and the large.

(Refer Slide Time: 35:07)



Another tRNA moves into the A site in order to add another amino acid to the peptide chain.

So, first the m-RNA comes to small subunit and it recognizes r-RNA sequence. Once it recognizes the bigger subunit of the ribosome comes to the place where the m-RNA is already bound. There is a sequence where it is anchored.

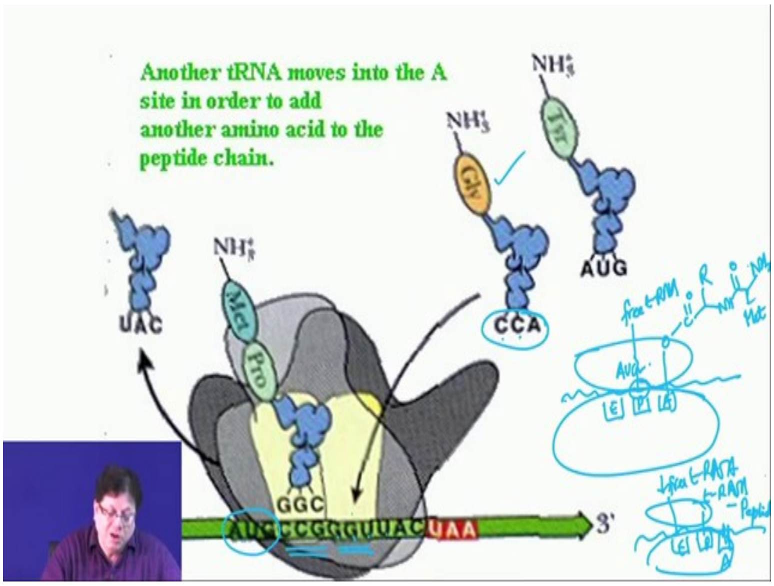
m-RNA has codons and then according to the codons tRNA will come and bind. Remember tRNA has anticodon. Now, this ribosome has different sites-P-site i.e. peptide site, A-site i.e. amino acid site, E-site i.e. is exit site.

tRNA will bind according to 3 letters codon and there will be an amino acid depending upon the codon present here. Another tRNA that will come and bind to it ok. So, this is your OCO one particular amino acid R1 NH_2 and this is your OCO, another amino acid is R1 and NH_2. So, this is the situation.

P site always has a consensus sequence i.e. AUG. What is AUG? AUG codes for methionine and the synthesis will start from there. So, this methionine containing tRNA will be the first one to bind to the P-site and then another amino acid will bind depending on the codon next to AUG. Now, there will be a reaction between the two. So, this reaction is nothing but a trans acylation. It is OCOR and then NH_2. Another tRNA will also have OCO and R1NH_2. Although they do not look proximity, actually they are quite proximal. Then this NH_2 comes, attacks here releasing the tRNA.

This is a free tRNA with OH here. So, you have OCO, then R 1. Then what you have is NH, then you have CO, and then you have R, and then you have NH_2 that means, you have now formed the first peptide bond.

(Refer Slide Time: 42:27)



Another tRNA moves into the A site in order to add another amino acid to the peptide chain.

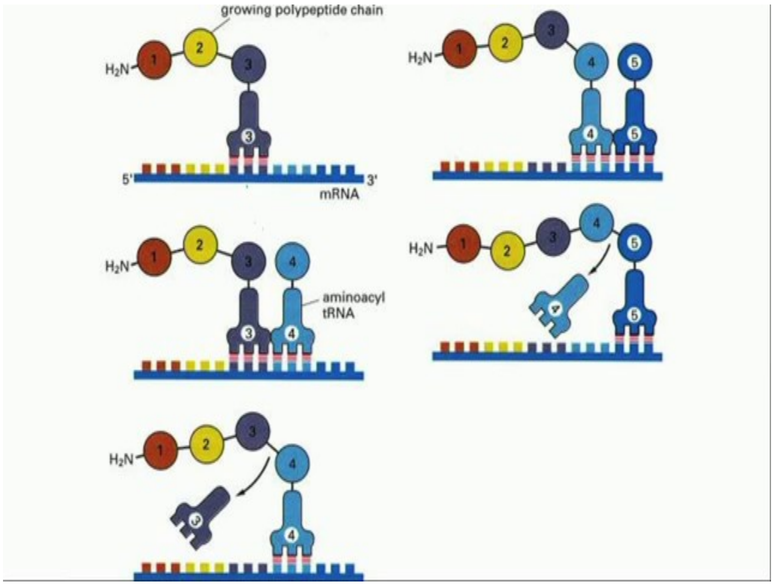
Here is your free tRNA because AUG which codes for methionine that is now transferred into the A-site. So, A-site is now having this dipeptide OCOR then NH then CO then this is methionine. I can write Met it will be always methionine and then NH_2 ok, so that is the scenario.

Free tRNAs will not now come to the E-site. P-site will hold the tRNA which is holding the dipeptide chain. The amino acid site is free.

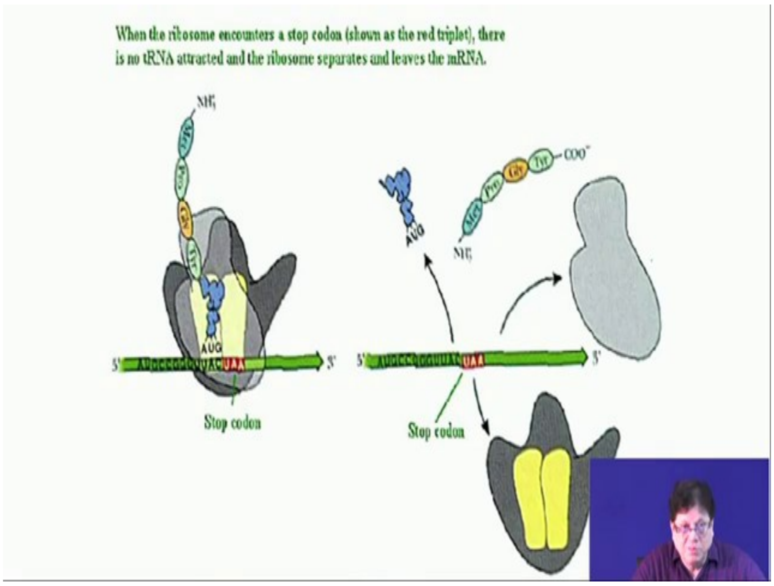
So, now, what will happen to another tRNA? Which tRNA will come that depend on the sequence of the m-RNA. So, another tRNA comes, and the same reaction happens and you will get a tripeptide. Then the ribosome slowly moves. So, this is the step where that tRNA this is exit, that means, the free tRNA now goes out. The growing peptide chain is transferred from P site to A site.

So, in this way the protein synthesis takes place. So, this will continue untill a stop codon comes. So, once a stop codon comes here, no tRNA will bind. Then the system knows that it is the end of the process. The whole protein is released.

(Refer Slide Time: 46:45)

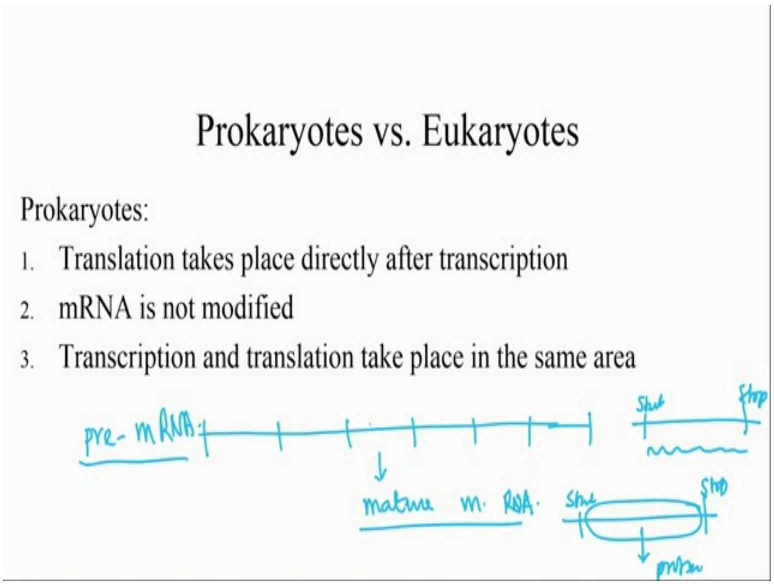


(Refer Slide Time: 47:15)



When the ribosome encounters a stop codon (shown as the red triplet), there is no tRNA attracted and the ribosome separates and leaves the mRNA.

(Refer Slide Time: 47:19)



Prokaryotes vs. Eukaryotes

Prokaryotes:

1. Translation takes place directly after transcription
2. mRNA is not modified
3. Transcription and translation take place in the same area

All proteins do not have methionine as the N-terminus. Many proteins have different N-terminus amino acid. There is another process which is beyond translation and that is called post translational modification. If methionine is not required as the N-terminal amino acid it will be removed by some enzyme and the correct protein will be made. It is called post translational modification.

Now, that is not for your eukaryotes. There is some problem. In prokaryotes, whatever messenger RNA is coming that is completely converted into the protein. There is no other processing needed in between this. But in eukaryotes, the m-RNA that is transcribed from the template strand of the DNA is called pre m-RNA. It is pre m-RNA because this m-RNA is made up of definitely all these codons. However, all the portions of it does not code for the protein.

In case of prokaryotes, the full mRNA will be made according to this codon sequence. But in case of eukaryotes, the pre mRNA will have to be further processed to make a mature mRNA. What is mature mRNA? The full portion from start to stop is transcribed. Then only this portion will be translated into the protein. This contains some portions which are not passing any information. This portion does not code for any amino acid. We will discuss it in the next lecture.