Cell and Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Week 10
Central Dogma of Molecular Biology (Part 2)
Lecture - 40
Translation

Hello everyone, this is Dr. Vishal Trivedi from the Department of Biosciences and Bioengineering, IIT Guwahati. So if you would like to understand the whole process of translations, what we have to understand is that we first have to understand the machinery. So machinery, which includes ribosomal RNA, tRNA, and messenger RNA, requires us to understand the different types of processes. So let's see what the different machinery is and what has been involved in the process of translation. So, as far as the machinery is concerned, we have different types of RNA molecules, such as ribosomal RNA, tRNA, and messenger RNA.

The ribosomal RNA is going to be responsible for the synthesis of the ribosome, and that ribosome is actually going to be the central machinery that is going to be responsible for the synthesis of proteins. Transfer RNA is actually going to transfer, or it is going to carry, the amino acids. So, it is going to carry the amino acids and transfer the amino acids according to the information given on the messenger RNA as well as according to the instructions from the ribosome. A messenger RNA actually contains the message, and that message is interpreted by the transfer RNA as well as the ribosomal RNA, and then through the ribosome, it gives you the protein.

So, we are going to understand the structure of some of these RNA molecules. So, let us start discussing the machinery. So, we are going to start by discussing first with the messenger RNA. So, as the name suggests, the messenger RNA, or mRNA, has a 5' prime end, 5' prime UTRs, and the ribosomal binding site. Then it also has the coding sequence, and it has the 3-prime UTR.

So what you see here is that the messenger RNA is going to have the 5-prime end, and this is the 3-prime end. So on the 5-prime end, it's going to have cap structures, and next to the cap, it is going to have the 5-prime UTRs. Then it is going to have the translational start site, and then it is going to have the translational stop site, and the nucleotides that are present between the start and the stop site are going to be called the coding sequence. So these are the coding sequences that are actually going to direct the protein synthesis machinery, like the ribosomes or the chi RNase, to synthesize the protein. And that is being done because it's actually going to contain the genetic code, and this genetic code is actually going to provide the information on what sequence you are going to assemble the

amino acids, and that's how it is actually going to give you the correct proteins.

And then we have, messenger RNA has three reading frames, out of which only one codes for the desired proteins. If in the sequence of these bases there is no stop codon to interpret the translation, then the synthesis of the entire polypeptide chain is called the open reading frame. So this is also called the coding sequence or the open reading frames. Then, next to that stop codon, you are going to have the three prime UTRs, and then you are also going to have the poly A tails. So these poly A tails, I think, remember when we were discussing transcription, we said that you require a cap and a poly A tail for the proper functioning of messenger RNA.

Both the poly A tail and the cap are actually required for messenger RNA to bind to the ribosome, and that's how they are going to participate in protein synthesis. Now, what it means is that when you actually have the coding sequence, the coding sequence can be interpreted into the three reading frames. So, let us take an example of how it is actually going to be done. So, for example, if you have a coding sequence like this, you have a coding sequence like A U G, G C G, A U G, C C C, A A A. Now, if this coding sequence can be interpreted in the three reading frames, the first reading frame can start from one, right? And you can have this reading frame; you can have it like this.

So this is the reading frame number one, okay? Now, if I change the color, right? So you can have reading frame number two, right? So reading frame number two is going to be like this: So it's going to leave the first residues and it's going to have the reading frame number two, which is UGC, then CCA, then UGC, then CCA, and A. So these are not going to be part of the reading frame because you only have two nucleotides. And this is all we are going to discuss in that. And then you're going to have the third reading frame. So that is also going to be there.

So you're going to have the third reading frame. So, what is going to be? First, we just started with this residue; right? The second one, we started with this residue, and the third one we are going to start with this residue. So, it's going to be like this, okay? When you see one A is left, this is what happened, right? This means that the three reading frames can actually start from different points: the first reading frame is going to start from A, the second reading frame is going to start from U, and the third reading frame is going to start from G. So, technically, a single RNA molecule or single messenger RNA molecule is actually going to potentially give you the three proteins, but you know that when we started with the reading frame number one, everything was settled. So, you are getting all the nucleotides; all these codons are being used.

So, that is actually going to give you a protein, whereas when you go with the rating

frame number two, what you see here is the first nucleotide you have left. And there are two nucleotides that are left at the back. So, there will be a truncated protein that is going to be synthesized, and as you can see, the nucleotides we are using for the synthesis are also going to be different. So, that is why the different types of proteins can be synthesized if you go with reading frame numbers 1, 2, and 3. Ideally, every gene is actually defined by which reading frame they are going to use.

So, that is how they have actually been able to provide the correct proteins or how they will be able to help in the synthesis of the correct proteins. This combination of triplets is actually called the genetic code because it provides information about which amino acid it codes for, and that is how it is referred to as the genetic code. So, let us discuss the genetic code. So, the genetic code is actually the messenger RNA having a random sequence of nucleotides differentiated by the bases attached to them, which are uracil, adenine, cytosine, and guanine, and the three nucleotides together code for a specific amino acid, which is called a codon. Now, the question arises: why are there only 3 nucleotides? Why not 4, or why not 2, right? So, if you go with the calculations, what you see here is that the 3 nucleotides actually give you a sufficient number of codons so that you can code for all 20 amino acids.

So, you know that we have 20 amino acids, right? You want to code for the 20 amino acids; the minimum number you require is actually 20 codons, right? So, 20 codons, how are you going to get the 20 codons, right? So, why are there only 3 nucleotides as a part of the genetic code? The genetic code is a triplet code called a codon. It is known that we have only four different types of nucleotides that make up the whole genome. So, it is known that each codon consists of 3 nucleotides, which means that 4 to the power of 3 is actually going to give you the 64 possible codons for the different types of amino acids. However, we have only twenty amino acids. So, it is obvious that more than one codon is coding for a single amino acid, and this is also illustrated by the wobble hypothesis.

So, if you go with the two nucleotides. Okay, then how many numbers are going to be 4 to the power of 2, right? So, what will the number be? It is going to be 16 nucleotides, right? So, if it is 16, it is actually lower than that number, right? And if you go with the 3 nucleotides, then the number is going to be 4 to the power of 3, right? And 4 to the power of 3 is actually 64. So, 64 is a sufficient number that can actually code for all the amino acids utilized by the ribosomal machinery for synthesis. So, what you see here is that this is a table of the genetic code. So, this is called the genetic code table, and you can use this table to know which amino acid is actually going to be incorporated.

So, what you can go with is that the first nucleotide is actually one of these. Then the second nucleotide and the third nucleotide, so for example, if I say U as the first, U as the

second, and U as the third, then what will the genetic code be? It is going to be UU and U, which is actually present here, right? So this is the first nucleotide; then we can have UUC, right? So U, U, and you can have C from this side, right? So, it could be phenylalanine. So, U U U is actually coding for phenylalanine. You can see that U U C is coding for phenylalanine, but the other amino acids and other combinations are actually coding for leucine. So, you can see that for the single amino acid, raphylaniline, you have multiple codons.

Similarly, you can see that for serine, you have the four nucleotides. Similarly, for tyrosine, you have two different types of nucleotides and so on. Apart from these, you can also have some of the stock codons like UAA, UAG, and U. GA and all these are called stop codons because they do not code for any of the amino acids. So, ideally, for the genetic codes, what you have is the 61 genetic codes that code for the amino acids and the 3 genetic codes that are actually called stop codons.

So, let us see what the general properties of the genetic code are. Genetic codes are triplets, so the genetic code is a triplet called codons; these are going to be triplets, like for example AUG, and it's always read in triplets. Okay, so you cannot have any other combinations; each coding sequence has a start and stop codon to initiate and terminate the translation, respectively. For example, the start codon is AUG, which codes for methionine, and the stop codons are UAA, UAG, and UGA. In some cases, the starting codons are GUG or UUC, so these are present in bacterial cases; in prokaryotic cases, you can also have GUG or UUC as a start codon.

So what is meant by the start codon is that whenever translation starts, the AUG will be the first codon present in the gene. The code is unambiguous, which suggests that the code is for only one amino acid, which means the code is unambiguous; it’s not like AUG is a codon for methionine in E. coli. But the AUG is a codon for lysine, leucine, or arginine in the other organism, so it's actually a constant codon; it does not vary from species to species or organism to organism.

That is called unambiguous. The AUG is fixed; AUG is fixed for the methionine, irrespective of these organisms. Then when you have the genetic code, there is no gap in the code, which means you cannot have the codon like this, or you cannot have the codon like A; then there is a gap, and then you can have AUG. So, this is going to be a problem with the codon. So, it is always going to be a continuous triplet, which is going to make the codons. The codon is degenerate; this means that one amino acid has more than one codon.

For example, phenylalanine is specific and is coded by the two codons UUU and UUC.

So the codons are degenerate, which means that for the single amino acids, you're going to have multiple codons. Right? You might have seen that we have 20 amino acids. And we have 61 different codons. So that's why ideally, for every amino acid, you can have either 2 or 3 amino acids.

Except there are exceptions that for the tryptophan and as well as for the methionine, they are coded by a single codon. Then the codon is non-overlapping; for example, a code like AUG, and then the codons will be AUG, CUC, GUG, and so on, and not like AUG, UGC, CCU, and so on. So, it is not like you can have the codons; for example, you have a sequence like this: AUG. It is not like you can have a codon like the first codon like this, then the second codon like this, or the third codon like this. So, you cannot have overlapping codons.

So, they are not overlapping. This means AUG will be like this; then the second codon is like this. So, they are not going to have overlapping nucleotides. The genetic code is universal, which suggests that the genetic code and its meaning are common to all life forms. However, there are some exceptions to this rule. For example, UGA is a stop codon, but it codes for tryptophan in mycoplasma, spiroplasma, and the mitochondria of some species.

That means that the genetic code is universal, which means if AUG is used for methionine, it will be used for methionine irrespective of the organism or the gene. But there are exceptions; for example, AUGA is a stop codon in all other species, but it codes for tryptophan in the case of some organisms, such as mycoplasma, spiroplasma, and some mitochondrial species. Similarly, you have another example like CUG. CUG codes for leucine in general, but in mitochondria, it codes for threonine. So, there are exceptions in biology, and that is why biology is so much more complicated and diversified.

Now, let us discuss the next machinery. So, the next machinery is the transfer RNA, and we have already said that transfer RNA is actually going to anchor the amino acid, and then it is actually going to deliver the amino acid, or it is going to transfer the amino acid during protein synthesis. As per the instruction from the ribosome, transfer RNA is a cloverleaf structure in two dimensions, and it has an L-shaped structure in three dimensions. It's going to have the L-shaped structure like this, something like this, in the 3D structure, and it's going to have the cloverleaf structure like this in a two-dimensional structure. The total length of the tRNA is 73 to 94 ribonucleotides, and it's going to have different shapes, so you're going to have the 5' end and the 3' end.

And, the tRNA molecule consists of the 5 prime phosphate terminal and an acceptor arm that ends in the CCA terminal at the 3 prime end. So, the 3 prime end is going to have the

CCA arm, which is also called the amino acid arm. So, this is called the CCA arm, and this is the arm that is actually going to receive the amino acid. Then, you also have the D-loop.

So, this is the D-loop. So, you are going to have the D-loop, which often contains some modified amino acids and nucleotides like dihydrouridine. Then, you also have the anticodon loop. So, anticodon loops are actually going to have anticodons. And the purpose of the anticodon is to interpret the codon. So, depending on the anticodon, the amino acids are going to be tagged to the tRNA.

So, the amino acid whatever is actually going to be tagged onto the 3 prime end with this CCA end; the anticodon is going to be present. So, the anticodon is actually going to check the codon, and then accordingly, it is going to bring that particular amino acid. Then you also have the T arm and/or the T Psi C, where Psi is called Pseudouridine. So you also have the T Psi C arms, and then you also have the extra arm. So this is the extra arm that is actually taking care of the extra nucleotides present in the tRNA.

So whatever extra nucleotides are present are being put into this extra arm. And the most important region of the tRNA is the two regions. One is called the CCA arm, which is responsible for the coupling of the amino acid, and the other is the anticodon arm, which is actually going to interpret the codon on the messenger RNA, and that's how these tRNAs are actually going to bring the amino acid for the coupling. Then we have the transfer RNA; each transfer RNA is specific to an amino acid, and it carries it on to the CCA arm. There are 30 to 45 different types of tRNA in prokaryotes and 50 different types of tRNA in eukaryotes.

For example, for glycine, there are two tRNAs that are specific: tRNA-Gly-1 and tRNA-Gly-2. So, remember that we have different types of codons. Accordingly, the tRNAs will also be different. So, we can also have multiple copies of the tRNAs, and these tRNAs are actually going to have two pieces of information.

First, they are going to carry the amino acids. So, they are going to have information about the amino acid. The second is that they are also going to have the information about the codon which they are going to, you know, use. So, that is what they are going to do with the help of the anticodon, and that is why we have 30 to 45 different types of tRNAs in prokaryotes. Whereas the number of tRNAs in the eukaryotic system is even larger, you have 50 different types of tRNA, which means that for a single amino acid, you can have multiple types of tRNA. One of the examples is that for glycine, it has two different types of tRNA molecules.

One is called Gly-1, the other is called tRNA Gly-2. Now, let us move on to the next machinery, which is called the ribosomes, or ribosome, or ribosomal RNA, which is coupled with the protein to give you the ribosomes. So, the ribosome is the ribonucleotide particle that contains the rRNA and the proteins; each ribosome is made up of two subunits: the large subunit and the small subunit. In prokaryotes, the mitochondria and the chloroplast contain 70S ribosomes, which are composed of the 50S and 30S subunits. In E. coli, the 30S subunit consists of the 16S ribosomal RNA and 21 ribosomal proteins, whereas the 50S subunit actually contains the 23S ribosomal RNA, 5S ribosomal RNA, and 31 different types of proteins.

In eukaryotes, it also contains the 87S ribosome. So, you can have two different types of ribosomes. You can have the 70S ribosome, which is actually present in prokaryotes, or it can also be found in the mitochondria or the chloroplasts. Then you can have the 80s ribosomes. These are going to be present in the eukaryotes. In eukaryotes, there are 80S ribosomes which consist of the 60S and 40S ribosomal subunits.

The 60S subunit consists of the 28S ribosomal RNA, the 5.8S ribosomal RNA, and the small 57S RNA, and it also contains approximately 50 different types of proteins. The 40S small subunit consists of the 18S ribosomal RNA and 33 ribosomal proteins. So, this is the composition of the ribosome where you can have the large subunit, the small subunit, and when these large and small subunits combine, they actually give you the three different types of activity sites. You can have the E site, which is also going to be called the exit site; you can have the P site, which is called the peptidyl site; and you can have the A site, which is called the A site or the amino site.

So, the function of these sites is very different. So, you can have the P site, which is for the peptidyl tRNA binding site, the A site, which is actually called the aminoacyl tRNA binding site, and the E site, from where the deacylated tRNA is going to exit. So, that is why you have the E site from where the tRNA is going to be removed. Then you have the P side where the tRNA is going to sit, right? So the P side, where you have the peptidyl tRNA that is going to bind, and the A side, where the incoming tRNA is going to come, right? So this means the ribosome is actually providing the necessary infrastructure and framework for protein synthesis; on one side, it is actually going to have the entry and exit of the tRNA, right? So tRNA is actually going to have the entry and exit onto the ribosomes, and on the other end, it is also going to have the messenger RNA binding site, so it's actually going to interpret the message that is present on the ribosome on the messenger RNA, and that's how it is actually going to be the central machinery where protein synthesis is going to take place. So, what have we discussed? We have discussed the structure of the translational machinery, we discussed the genetic codes, and now we are going to start discussing the mechanism of translation. And the first step in the

mechanism of translation is that you want to activate the amino acid, which means you are actually going to couple the amino acid to the tRNA.

That is actually how you are going to couple it to the tRNA. As soon as you couple it to the tRNA and make the tRNA aminoacyl or tRNA acylated, it will be committed to protein synthesis. So, let us see how we can perform the amino acid activations. So, during this process of amino acid activation, the amino acids are attached to the tRNA in the presence of an enzyme called aminoacyl tRNA synthetase. This enzyme activates the amino acid by covalently attaching to the tRNA.

When the tRNA gets charged, it is called aminoacyl tRNA. So, during this process, the amino acids are attached to the tRNA with a high-energy bond, so they are actually called activated amino acids. So, what will happen is that you have an amino acid, and then you have a tRNA, and then you are actually going to utilize the energy in the form of ATP. So, what happens is that the aminoacyl tRNA synthetase is actually going to utilize the ATP, and it is going to convert the ATP into AMP and synthesize pyrophosphate. Breaking the ATP, it is actually going to generate energy, and that energy is going to be utilized for the formation of the coupling of the amino acid to the tRNA, and that is how you are going to have the aminoacyl-tRNA. This aminoacyl tRNA is also called the activated amino acid, and that activated amino acid is going to participate in protein synthesis.

Once the amino acid activation is over, it is going to participate in the translations. You know that we have already discussed that we have the initiation codon, or we have the first codon that is actually going to be used as the initiation. So, the question is what the difference is: will there be a difference between the tRNA that is going to be responsible for the initiation of the translational machinery, or is it actually the regular tRNAs? So, in the eubacteria, the first amino acid in the polypeptide chain is N-formylmethionine, which is specific to the three codons AUG, GUG, and UUC. Remember that we said that AUG, GUG, and UCG can be potential initiation codons. These initiation codons are not coding for the normal amino acids; they are not coding for the normal methionine, but for the N-formylmethionine.

And that is why for the first tRNA that is going to participate in the initiation reactions, it is not the methionine that is going to be tagged to that particular tRNA; it is the n-formylmethionine that is going to be used. So, in this process, what will happen is that methionine is actually going to be formulated to generate the FMET or the n-formylmethionine, and then this is actually going to be coupled onto tRNA to generate the FMET tRNA. And then that is how it is actually going to generate the FMET tRNA, and that FMET tRNA is the tRNA species that is going to participate in the So, wherever

these codes are present at the initiation point, they code for the n-formyl methionine, but when they are present in between the coding sequences, they code for methionine and valine, respectively. Which means if the AOG is present at the first codon, it is actually going to code for the n-formyl methionine, but if the AOG is present inside the codon, inside the sequence, then it is actually going to code for the normal methionine. So, how does this happen? This happens because of the difference between the initiator tRNA and the one you used during the process of translation.

So, initiator tRNA has a unique feature that distinguishes it from the elongating tRNA in bacteria. So, you can have two different types of tRNA molecules. The tRNA that is responsible for the initiator tRNA, and then you can also have the tRNAs that are participating in the elongation steps. Now, let us talk about the first step, and the first step is called the initiation. So, in the first step, the small subunit of the ribosome binds to the messenger RNA such that the initiation codon lies in the partial pea size.

This is made possible by the activity of initiation factor 3. So, it basically prevents the untimely reassociation of the large and small subunits of the ribosomes. Moreover, it promises the accuracy of the initiation sites. So, how is that being done? In the messenger RNA, there is a ribosomal binding site, which consists of the Shine-Dalgarno sequences and the initiation codon, right? So, this shinedergano sequence, which is the pi prime AGG AGG U, is located tens of base pairs upstream of the initiation codon and is complementary to the region near the 3 prime end of the 16 S ribosomal RNA, a component of the small subunit. So, what happens is that you have the shine-dalgarno sequences which are present on the messenger RNA, and then these shine-dalgarno sequences actually have affinity for the small subunit of the ribosome. And when these two come together, the Scheindl-Gonus sequence actually helps to align the two RNA species in such a way that the initiation codon is actually present in the T site.

So, this is what happened: you have the initiation codon, which is the AUG, and before this, you are actually going to have the Shine-Dalgarno sequence. So, in the next step, your tRNA carrying the n-formyl methionine enters the P-site and binds to the messenger RNA via its anticodon, right? Once the initiation codon 3 and initiation factor 1 are actually present, they will block both the E and A sites. So, you have the three sides where you have the E, P, and A. On the P side, you have the starting codon, whereas on the E side, as well as the A side, you are going to have the initiation factors 3 and 1. Then what will happen is that the initiation tRNA is going to come along with the initiation factor, which is actually a GTPase, and because of this, the initiation codon is actually going to bind to the P site with the help of the codon and anticodon interaction.

So, you see these are the codons, AUG, which actually code for N-formylmethionine,

whereas this is the anticodon for N-formylmethionine. So, you are actually going to have the codon and anticodon, which are going to be unique for this particular sequence. So, you see that it is actually complementary to one another, right? So, G, C, U, A, and AU. So, this is why it is actually going to be recognized by this particular initiation codon, okay.

So, initiation factor 2 is responsible for this activity. It directs the initiator tRNA to its correct position in the initiation complex. It also exhibits ribosome-dependent GTPase activity. Once the GTP is hydrolyzed, the 57-rubinate joins to form the complete ribosome. Finally, when the large subunit also joins the complex, it forms the complete P side and A side, and the second charged tRNA enters the A side. This tRNA, as per the rule, has the anticodon corresponding to the codons in the messenger RNA.

So, what happens is that once this step is over, you are actually going to have the entry of the 50S ribosome, and then the initiation factors 3 and 1 are actually going to be released along with initiation factor 2 and GTP. It is actually going to allow the binding of the large subunit, and that is why it is going to assemble the complete ribosomal machine. So, this actually means that the initiation is going to be over in the case of prokaryotes. So, there are differences between the initiation into prokaryotes and eukaryotes as well. So, all other steps are almost identical except that the initiation steps are different between prokaryotes and eukaryotes.

In the translation process, the main difference between eukaryotes and prokaryotes is in the initiation process itself. Some major differences between the eukaryotic initiation and the prokaryotic initiations are as follows. In eukaryotes, there is only one start codon, AUG, which codes for methionine and not n-formyl methionine, whereas eukaryotic cells need more initiation factors than prokaryotes. So, there are two major differences: the eukaryotic AUG code for methionine, whereas the prokaryotic AUG code for n-formylmethionine, and the eukaryotic cell needs more initiation factors than the prokaryotes. Eukaryotic cells require the 12 initiation factors, whereas prokaryotic cells require only the 3 initiation factors, initiation factors 1, 2, and 3.

In eukaryotes, the process of association of the messenger RNA with the small subunit, the 40S ribosomal subunit, is more complex than in prokaryotes. The 40S ribosomal subunit identifies the 5' methylated cap of messenger RNA, and then there is a scanning process involved in which the initiation codon is recognized. This recognition is added by the ATP-dependent helicase that hydrolyzes ATP. This recognition of the initiation codon is also being aided by the COSAC sequences. And these COSAC sequences are almost similar to the Shine-Dalgarno sequences which are present in prokaryotes.

So, this is one of the major differences that you have a scanning step in the case of the eukaryotic system, which means you are actually going to have the messenger RNA, and this messenger RNA is actually going to be scanned by the small subunit for looking for the AUG sequence. This scanning is being aided by the COSAC sequences because the COSAC sequences that are present on the ribosomal small subunit are actually going to help the positioning of the small subunits. So, these are some of the classical differences between the initiation of messenger RNA and the initiation of translation between prokaryotes and eukaryotes. Then we step, then we enter into the elongation steps.

So, elongation steps are actually a cyclic process. The elongation process starts from the formation of the first polypeptide band to the addition of the last amino acids. The amino acids are added to the nascent polypeptide chain one at a time. Addition of the amino acid is a very rapid process. The peptide sequence is in the order of codons and anticodons in the messenger RNA.

The rate of elongation is nearly 15 amino acids per second. There are three more requirements regarding the elongation. So, what are the requirements? You also require the messenger RNA and the 70S ribosomes; you require the aminoacyl tRNA for all the amino acids that are actually going to be present on the messenger RNA in the form of codons, and then you also require the elongation factor. So, what elongation factors are required? You require the elongation factor Tu, the elongation factor Ts, and the elongation factor G. So, elongation-factor TU is a G protein that actually binds to the aminoacyl-tRNA and directs it to the correct position at the ribosome A site. Whereas the elongated receptor Ts has as its main function to generate the EFTU, regenerate the EFTU, and hydrolyze GTP, the EFG is also a G protein that mediates the translocations.

So, elongation has multiple steps; it requires this type of machinery, the messenger RNA, and the complete ribosome; it requires the aminoacyl tRNA, which is actually going to carry the amino acid, and they also require the elongation factors. So, elongation is carried out by the ribosome in three stages. One is decoding; the second is peptide bond formation, and the third is translocation. So, decoding is being done by the interaction of the codon, and I think we have already seen how the initiation codon has the anticodon for the codon onto the AUG. So, it is codon-directed binding during the process of ribosomes selecting and binding to the incoming aminoacyl tRNA at a site whose anticodon is complementary to the codon of the messenger RNA.

Decoding region of the 16S ribosome confirms the proper base pairing between the codon as well as the anticodon. So you are going to have the codon, and then you are going to have, for example, if you have the codon AUG, it is going to have the anticodon C. So, you are going to have the anticodon, and that is going to make the base pairing

according to the hydrogen bonding, and that is how it is actually going to create the firm interaction between the tRNA and the messenger RNA that is present on the codon in the messenger RNA. Once this is done, you are going to have the second step, which is the peptide bond formation. So, in this process, the peptidyl group of the P site of tRNA is transferred onto the aminoacyl group in the A site to form the peptide bond.

So, in this step, what will happen is that in the ribosomes you are going to have the aminoacyl tRNA, the initiator amino acid, and that is present on the P site. So, you are going to have the amino acid that is present on the tRNA. What is being presented on the P side? So, from here it is actually going to be coupled onto this, and that is how this amino acid is actually going to have the amino acid; and then it is going to have the initiation amino acid, and that is how this is going to continue. Then there will be a translocation. In the translocations, the tRNA of the A site is transferred to the B site to make space for the next aminoacyl tRNA at the A site, and the A site of tRNA is shifted to the E site.

This shift is also coupled with the ribosomal movement along the messenger RNA. So, what is meant by translocation? Translocation means that this tRNA is going to move on to this, and this deacylated tRNA is going to move into the E site, and that is how it is actually going to be removed from the ribosome. So, let us see how this is done. So, in the process of chain elongation on the ribosome, EF-Tu promotes the entry of the aminoacyl tRNA into the A site of the ribosomal RNA. First, the EFTU binds to GTP and activates the EFTU-GTP complex, which binds to the tRNA. When the codon and anticodon base pairing is stabilized, the hydrolysis of GTP occurs, converting GTP into NPI, which helps in the binding of aminoacyl-tRNA to the A site, and after this, EFTU is released.

So, this is what is going to happen. The EFTU is actually going to combine with the GTP. So, that is how it is actually going to have the binary complex. Binary complex means the two molecules are coming together, and once the binary complex is formed, it is actually going to bind the tRNA, which is going to be different from the initiation tRNA. This tRNA is going to contain, you know, the amino acids, and this is going to have the anticodon. So once this anticodon is matched onto the codon on the messenger RNA, then this GTP is going to be hydrolyzed, and this GTP is going to be converted into GDP plus PI, which means And, once this happens, the EFTU is actually going to be released, and it is going to come back again. Its GDP is going to be released, and it is going to be replaced with another round of GTP, and that is how this cycle is going to continue.

So, that is how it will help bind the tRNA into the A site. Once the tRNA binds, there

will be a peptidyl transfer. So it is the peptide bond formation step in which the amino acids of the peptide bonds are linked to the tRNA molecule in the A site, and the carboxyl end of the peptide bond is uncoupled from the tRNA molecule into the P site. So the peptide bond is going to be formed between the amino acid that is present in the P site and the amino acid that is present in the A site. From this, this amino acid is actually going to be transferred onto this, right, and it is going to form the peptide bond.

So, that is why, if this is the A1, then it is going to form the A2, A1 like this. So, how is this being done? So, this reaction is carried out by the enzyme called peptidyl transferase. Peptidyl transferase is an enzyme that is associated with the 23S ribosomal RNA of the 50S ribosomal subunit, and the peptide bond formation involves the O to N migration and the conversion of an ester into an amide bond. So, this is what is going to happen: you have the FMET, which is the initiation amino acid, which is going to be present on the P side, and then it has the free carboxyl group. So, you have a free carboxyl group, and then there will be an attack from the amino group of the second amino acid, which is present on the A side, and that is how it is going to attack this carboxyl group, and that is why this reaction is going to be catalyzed by the peptidyl transferase. And then there will be an amide linkage that is going to be formed between the amino acid that is present on the A side.

So, you have this amino group, right? And this is going to be the carboxyl group, and there will be an amide linkage that is going to be formed. And then this amino acid that is present on the P side is going to be transferred onto this. And you still have the deacylated tRNA that is present on the P side.

Then we have the translocations. So, translocation means you are actually going to move the molecules. So, you are going to move the ribosomes onto the messenger RNA for one more codon. There are three things that are going to be necessary for the translocations. Deacylated RNA moves from the P-side. Peptidyl transferase moves from the A to the P site, and the ribosome should move on to the messenger RNA by one more codon so that the next codon can come to the A site. So, this means if you have codons like AUG, GG, and GCG, right? So, initially the ribosome is going to be present on the AUG, now it is going to move to the stop codon and it is going to be present on this.

So, this codon is now going to be present in the B site, and that is why another amino acid is going to enter the A site, while the tRNA that was initially present in the A site is going to be moved to the B site. And that the deacylated tRNA is what was present initially on the B site is actually going to be present on the E site, and from this site, it is actually going to be released. The translocation step is carried out by the EFG factors, elongation factor G. During the translocation, the acceptor ends of both tRNAs in the A

and P sites interact with the peptidyl transferase center of the 23S ribosomal RNA of the 50S subunit.

In translocation, tRNA A and B transfer to the P and E sites, respectively. As ribosomes move 3 nucleotides along the messenger RNA chain in the 5' to 3' direction, during the translocation step, GTP is converted into GDP and the uncharged tRNA is released from the P site to the E site. And the newly formed peptidyl tRNA is going to move from the tRNA on the A site to the P site, and the elongation process is nearly the same in both prokaryotes and eukaryotes. So this is exactly what is going to happen, right? You have the, uh, now what we have is a site where we have the tRNA, and that tRNA is going to contain the two amino acids a1 and a2, right? And on this side, you have the tRNA, which actually does not contain any of the amino acids. Because this amino acid is already being transferred onto another tRNA, and then you also have the mte side, so when this ribosome is going to shift in this direction. One codon right, then what will happen is that this is going to move to here, and that is how you are going to have the deacylated tRNA, so this is what is going to happen, right? So once this ribosome moves for one nucleotide, it will move here, right? This will be present here, and since this does not contain any of the amino acids, it will be released from the ribosome.

The same is true for this one, so this is going to be moved to this side, right? It is actually going to have the A1 and A2 on this. And when this happens, this site is going to take up the new codon, and that is going to have the new tRNA, and that is going to have the fresh amino acid according to the next codon, and this process is going to continue, which means the P is going to have the tRNA, which is going to have A1 and A2, and A is actually going to have the new amino acid. So, it is going to have the tRNA, which is going to have the A3, and then again the same process is going to happen: the A1 and A2 are going to be transferred onto the A3, and E is going to have the dose that tRNA, which, you know, does not contain the amino acid. So, this process is going to continue; first, you are going to have the coupling reactions, then you are going to have the translocation, and that is how this will continue, right? So, these elongation steps continue until they reach the stop codon.

So, what happened at the stop codon? So, then it is going to have the terminations. So, the termination of translation occurs due to the stop codon. So, there are three stop codons: UAA, UAG, and UGA that are present. So, out of these three, one of the stop codons appears on the A side of the ribosome, causing termination because there is no tRNA present corresponding to these codons. So, tRNA is not going to bind the codon and causes the termination.

So, during the termination, there is a release factor which are involved. So, when the

UAA or UGA is in the A site, RF1 binds to the ribosome, and when the UAA or UGA is in the A site, RF2 binds to the ribosome, and RF3 is a type of GTPase that has the main function of catalyzing the release process through GTP binding and hydrolysis. So, this is exactly what is going to happen, right? During the termination step, what will happen is that you are actually going to have the fully synthesized protein, which is going to be present on this side. So, you are going to have A1, A2, A3, AXX, and so on. But on this side, since this was moved one step further, you are going to reach the stock codon.

So, the stock codon is actually going to allow the entry of the tRNA, but this tRNA will not have any amino acid. So, it is not going to have the amino acid, and because of that, this is actually going to move in this direction, and it is actually going to be released, right? Because there is no amino acid on this side, it will not be able to do the coupling reaction, and that is why it is actually going to be released from the ribosomes. The release factors Rf1 or the release factor Rf2 bind to the ribosome nearly at the A site, and the polypeptide chains are released from the ribosome by the peptidyl transferase complex. The peptidyl transferase complex transfers the carboxyl terminal residue of the peptide chain from the tRNA in the PSI to the water molecule. Now, the release factor RF and the GTP are released, and the tRNA is also free from the ribosomes. And now, the 70th ribosome is unstable due to the presence of the initiation factors IF 1 and IF 2, and the ribosome recycling factors.

As a result, the 70th ribosome disrupts into the 30th and the 50th subunits, and it is going to be prepared for the initiation. So, this is exactly what is going to happen once it reaches this top codon. The GTP and the RF factors are going to come, and that is why they are actually going to release the newly formed polypeptide chain, as well as the 30S and 50S subunits.

So, this is all about the translations. What we have discussed so far is the charging of the amino acids. So we have discussed how the amino acids are going to be activated with the help of ATP, and that's how it is going to couple with the tRNA. So first, you are going to have the charging steps, right, where you are actually going to couple the tRNA with the corresponding amino acids. So it's going to form the aminoacyl-tRNA. And once this is synthesized, it will actually deliver the amino acids. So, where the tRNA, aminoacyl tRNA, is going to bind, if it is the initiation codon, then it is actually going to go and bind onto the P site.

And then, if it is the normal tRNA, it is actually going to go and bind to the amino acid, like the A site. And once it binds to the A side, the tRNA that is present, the initiation tRNA, is going to be, there will be a peptide bond that is going to be formed onto the tRNA that is present on the P side. To the tRNA, what is present on the A site will

continue this cycle. Whereas the tRNA, when it is getting deacylated, is going to be transferred to the E site, and from there, the tRNA will be released into the cytosol, after which it will again participate in the charging reactions. Whereas during the translocation, the tRNA that is actually going to have the dipeptide is moved to the P site, and then another tRNA comes to the A site. And that is how this cycle is going to continue until the ribosome reaches the stop codon, and at the stop codon, the release factors will come, and that is how they will actually be released.

The polypeptide is also going to disintegrate the small subunit as well as the large subunit. Because at the stock codon there will be no tRNA that is actually going to have the amino acid bound. So, they are actually going to activate, you know, allow the release factor to free the machinery from the messenger RNA. So, that is how the messenger RNA, as well as the small and large subunits of the ribosome, is going to be dismantled from the protein; this is machinery.

So this is all about the translation that we have discussed. Now we are going to discuss the post-translational modifications. So when we talk about post-translational modifications, the protein can be synthesized as a native protein, and then it can actually be converted into different types of modifications. These modifications could be the covalent modifications, the non-covalent modifications, the reversible modifications, or irreversible modifications. So, one of the classical examples of the reversible post-translational modification is phosphorylation, and the irreversible post-translational modification is ubiquitination. So let's discuss the post-translational modifications.

So, post-translational modification, the first post-translational modification we want to discuss is phosphorylation. So phosphorylation is actually done by the ATP. So when you have a protein, the protein is actually going to take up the phosphate from the ATP, and that's why it is actually going to form the protein phosphate, right? And you know that the phosphate is actually going to have the high-energy bond, right? So it's actually going to provide energy to the protein molecules, and that's why it is actually going to induce different types of modifications. It is actually going to make the modifications; it's going to have the structural modifications, so it can have the activity modifications, like it's going to increase or decrease the activity of the proteins. And it can also actually have the modification in terms of the association of the proteins with another protein or the substrate as well. So, when you have phosphorylation, it is actually going to induce multiple types of changes in the protein, and these changes could be advantageous or disadvantageous.

For example, sometimes it may actually increase the activity, and sometimes it is actually going to decrease activity. So, how is the phosphorylation going to happen?

Phosphorylation is always done on the amino acid. So, you have the classical three amino acids on which phosphorylation can be done. And you can also have the tyrosine. So all these amino acids actually contain a free hydroxyl group, and to this free hydroxyl group, you can actually attach the phosphate.

So, how did this happen? For example, I have shown you one reaction where I used serine. You have the OH, which is actually going to contain the lone pair electrons, and these lone pair electrons are attacking the phosphate molecule, and that is why, with these high-energy intermediates, This phosphate is going to be transferred onto the phosphate, onto the serine, and that is why it is actually going to form the phosphoserine molecules. So, you require some cofactors like magnesium, and this reaction is always catalyzed by an enzyme called kinase. So, in this case, if this is the serine that is getting phosphorylated, then this can be called the serine kinase. Uh, how the phosphorylation— I think we have already discussed that the phosphorylation is actually going to cause conformational changes in the protein. So phosphorylation causes the conformational changes in the phosphorylated protein; these conformational changes stimulate the catalytic activity of the protein, so the moni protein can be activated or inactivated by phosphorylation.

The phosphorylated proteins employ the neighboring protein that has the structurally conserved domain that distinguishes and binds to the phosphomotor. These domains are specific to the diverse amino acids. Protein phosphorylation is a reversible post-translational modification that is carried out by kinases. which phosphorylates and the phosphatase which dephosphorylates the substrate.

There are two types of enzymes that make possible the dynamic nature of the phosphorylated protein. So, the balanced concentration of the kinase and the phosphatase is very important for the cell, and it is important for the catalytic activation of the particular phosphorylation site. So, you can imagine that you have two different types of proteins. The protein can be present as the P, or it can be present like this. So, if when you have the So, you can have the active enzyme right. So, when you add the help of the kinase, you can take the ATP and actually add the phosphate; then you are actually going to form the phosphorylated enzyme.

So, it could be an active enzyme once the signal is over. So, if you once get the signal from some kind of stress or some kind of ligand. So, for example, you can have the insulin, right? Insulin actually goes and binds to the insulin receptor, then it constitutes a signal; once the signal is there, that signal is going to be perceived by the protein kinase. And that protein kinase is going to say, "Okay, phosphorylate this particular protein," and once there is phosphorylation of this protein, it is actually going to activate that protein.

But once the signal is over, which means once you have removed the insulin from the vicinity, the signal is going to die, and once the signal is over, that is actually going to activate the corresponding phosphatase. And that corresponding phosphatase is actually going to remove the phosphate molecule from the protein, and that is how it is actually going to generate the inactive enzyme or revert to the native enzyme. So this is one of the classical examples of how phosphorylation can actually be used as a tool to modulate protein activity, and that's how you can respond to the external environment.

Now, apart from phosphorylation, you can also have glycosylation. So, glycosylation is a direct function of the biosynthetic secretory pathway in the endoplasmic reticulum and the Golgi apparatus. Approximately 50 percent of the protein characteristically expressed in a cell grows through this alteration, which involves the covalent addition of sugar moieties to specific amino acids. Mostly soluble and membrane-bound proteins expressed in the endoplasmic reticulum undergo glycosylation, including all secreted proteins, surface receptors, and ligands. Moreover, some proteins that are transferred from the Golgi to the cytoplasm are also glycosylated.

So glycosylation is a topic that we are actually going to discuss in detail when we talk about vesicular transport. So, I am not going to discuss the glycosylation mechanism, but what is meant by glycosylation is that the protein is going to be connected with the carbohydrate molecules or a combination of the carbohydrate molecules. And that is why it is actually going to form the glycoprotein, and this glycoprotein is actually going to have different types of sugar molecules attached. Based on these sugar molecules, it is actually going to provide the signal to the cell, and that is how these protein molecules are going to be distributed to the different droplets. So, that is what we are actually going to discuss in detail when we talk about vesicular transport.

Apart from this, we can also have ubiquitination. So, ubiquitinylation is a covalent modification of the protein, and it is an irreversible modification. So, this is another pro-translational modification where ubiquitin, which is a protein, is added to another protein. So, ubiquitin is the eukaryotic protein coded by four different genes in mammalian cells, such as UBA52, RPA27, AUB, and UBC. The ubiquitin is made up of 76 amino acids and has a molecular weight of 8.5 kDa; it is characterized by the presence of a C-terminal tail and seven lysine residues in ubiquitin relations, basically the carboxyl acid of the terminal glycine.

The diglycine motif in the activated ubiquitin forms an amide bond to the epsilon amino group of the lysine in the modified protein. It marks the cellular protein for the process of degradation by the proteasome, changes the protein locations, and prevents or promotes the protein-protein interaction. So, ubiquitinization is required for protein degradation.

So, when the protein is actually going through the aging process and is non-functional, it will undergo ubiquitinization. And this ubiquitinization is also going to sometimes alter its localization or sometimes alter the protein-protein interaction, which means if you actually mask some of these sites that are responsible for the protein-protein interactions with ubiquitinization, you will also be able to disrupt the protein-protein interaction as well.

So, what are the different steps that are involved in ubiquitin annihilation? In step one, you are going to activate the ubiquitin. So, it occurs in a two-step reaction process. At first, the ubiquitin interacts with the ATP and forms the annihilated ubiquitin or ubiquitin annihilated intermediate. In the second step, ubiquitin is transferred to the E1 active site containing the cysteine residue. So, you can actually have the activation of the ubiquitin. So that is the first step, right, where you are actually going to have the ubiquitin, it is going to attach to the ATP and it is actually going to form the adenylated ubiquitin, and that adenylated ubiquitin is actually going to couple onto the ubiquitin enzyme, right, the E1 enzyme, with the help of the sulfur.

So this is going to form. So E1 is actually going to contain the cysteine, and that's how it is going to utilize that cysteine to couple the ubiquitin. Then, we have the ubiquitin E2 enzyme. So, the transfer of ubiquitin is going to be done from the E1 side to the E2 side via a trans-esterification reaction. So, from the E1, the ubiquitin is going to be transferred to the E2 by a trans-sterilization reaction. Now, this E2, as well as the substrate protein, is actually going to bind to the E3. And, in the last step, the ubiquitin cascade, there is a formation of an isopeptide bond between the lysine of the target protein and the C-terminal glycine of the ubiquitin via the activity of one of the hundreds of E3 ubiquitin ligases.

So, this is going to be E3, which is going to be called E3B ubiquitin ligase. And it is actually going to couple the ubiquitin that is present on the E2 to the substrate protein, which is actually going to have the free amino group, and that is why the ubiquitin is going to be transferred to this free amino group. And you can imagine that this process, if you continue for several rounds, will actually lead to multiple ubiquitins being transferred onto the substrate molecule. So, on the substrate molecule, you can have mono-ubiquitinylations or polyubiquitinylations, and depending on the mono or polyubiquitinylations, the protein's properties can be modulated. Apart from that, we can also have other kinds of post-translational modifications that are called methylation.

So, this process refers to the addition of a methyl group to the nitrogen, the oxygen, or the amino acid side. The N-methylation is irreversible, whereas the O-methylation is potentially reversible. Methylation increases the hydrophobicity of the amino acid and

neutralizes the negative charge when attached to the carboxylate. The main methyl group transfer for such a reaction is SAM or S-adenosylmethionine. This reaction is mediated by the enzyme methyltransferase.

The methylation process is involved in epigenetic regulation, such as histone methylation and demethylation. So, methylation is a process that is actually responsible for the opening and closing of the chromatins, and that is how it is going to be responsible for the activation or deactivation of the transcription processes. So, it is a kind of mechanism that is used for regulatory purposes. So, for example, you know that there are heterochromatin and euchromatin, and the chromatin can be fractionated into heterochromatin and euchromatin. Methylation and demethylation can be processes through which you can actually convert heterochromatin into euchromatin and euchromatin into heterochromatin. So, with this, I would like to conclude my lecture here. Thank you.