



Google Analytics

Visits	Bounce Rate
0.09%	43.64%
0.19%	43.64% (0.00%)
27.27%	43.55%
85.19%	43.55%
56.52%	43.07%
95.45%	39.13%
92.31%	40.91%
85.71%	38.46%
100.00%	28.57%
40.00%	16.67%
0.00%	0.00%
	80.00%

Google Analytics

Visits	Bounce Rate
0.09%	43.64%
0.19%	43.64% (0.00%)
27.27%	43.55%
85.19%	43.55%
56.52%	43.07%
95.45%	39.13%
92.31%	40.91%
85.71%	38.46%
100.00%	28.57%
40.00%	16.67%
0.00%	0.00%
	80.00%

Google Analytics

MOLECULAR BIOLOGY

Prof. Vishal Trivedi
Biotechnology and Bioengineering
IIT Guwahati



INDEX

<u>S.NO</u>	<u>TOPICS</u>	<u>PAGE.NO</u>
Week 1		
1	Lec 1: Cellular Structure (Prokaryotic cells)	4
2	Lec 2: Cellular Structure (Eukaryotic cells)	16
3	Lec 3: Cellular Structure (Eukaryotic cells)	25
4	Lec 4: Cell Fractionation (Part 1)	36
5	Lec 5: Cell Fractionation (Part 2)	46
Week 2		
6	Lec 6: Cellular Metabolism (Part 1)	58
7	Lec 7: Cellular Metabolism (Part 2)	72
8	Lec 8: Cell Cycle and Control (Part 1)	88
9	Lec 9: Cell Cycle and Control (Part 2)	103
10	Lec 10: Program Cell Death	116
Week 3		
11	Lec 11: Biomolecules (Part 1: DNA)	132
12	Lec 12: Biomolecules (Part 2: DNA Sequencing)	150
13	Lec 13: Biomolecules (Part 2: RNA)	156
14	Lec 14: Amino acids	173
15	Lec 15: Protein	185
16	Lec 16: Enzymes	199
Week 4		
17	Lec 17: Genetic Material (Part 1)	213
18	Lec 18: Genetic Material (Part 2)	228
19	Lec 19: Genetic Material (Part 3)	239
20	Lec 20: Central Dogma of Molecular Biology	251

Week 5

21	Lec 21: Replication (Part 1: Prokaryotic System)	262
22	Lec 22: Replication (Part 2: Prokaryotic System)	279
23	Lec 23: Replication (Part 2: Eukaryotic System)	286
24	Lec 24: Mutagenesis and repair Mechanism	300

Week 6

25	Lec 25: Transcription in Prokaryotic system	313
26	Lec 26: Transcription in Eukaryotic System	325
27	Lec 27: Post Transcriptional modifications	333
28	Lec 28: Gene Control Mechanism (Part 1)	344
29	Lec 29: Gene Control Mechanism (Part 2)	360

Week 7

30	Lec 30: Translation in Prokaryotic system	368
31	Lec 31: Translation in Eukaryotic System	384
32	Lec 32: Post Translational modifications	396

Week 8

33	Lec 33: Southern Blotting	411
34	Lec 34: Northern Blotting	421
35	Lec 35: Western Blotting (Part 1)	433
36	Lec 36: Western Blotting (Part 2)	445

Week 9

37	Lec 37: Polymerase Chain Reaction (Part 1)	455
38	Lec 38: Polymerase Chain Reaction (Part 2)	472
39	Lec 39: Real-Time PCR	485

Week 10

40	Lec 40: Cloning (Part 1)	502
41	Lec 41: Cloning (Part 2)	511
42	Lec 42: Cloning Vectors	523

43	Lec 43: DNA Delivery (Part 1)	528
44	Lec 44: DNA Delivery (Part 2)	538
45	Lec 45: Screening of Recombinant Clones	545
46	Lec 46: Protein Over-expression	558
Week 11		
47	Lec 47: Genome Editing (Part 1)	567
48	Lec 48: Genome Editing (Part 2)	581
Week 12		
49	Lec 49: Applications of Molecular Biology (Part 1)	593
50	Lec 50: Applications of Molecular Biology (Part 2)	602
51	Lec 51: Applications of Molecular Biology (Part 3)	609
52	Lec 50: Applications of Molecular Biology (Part 4)	618

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati

Module - 01
Basics of Biological System (Part 1)
Lecture - 01
Cellular Structure (Prokaryotic cells)

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati. And in this module, we are discussing about the different types of cells. So, in today's class, we are going to start discussing about this particular cell and how the cell is actually functioning the different types of you know, task and how what are the different organelles are present within the cell and what is the structure of the cell. So, when we talk about the cell. So, as you can see that the higher you carry yours, I have multiple organs to perform the specific functions such as liver, kidney and heart, whereas in the and these organs have the specific tissue and each tissue is composed of the cells.

So, whatever is function what is happening in a higher organisms like the humans that it has the different types of organs like we suppose for example, we have the liver, we have the kidney, we have heart, we have the lungs and all these organs have their specialized functions, but these organs are also made up of tissues and that tissues are made up of the cell. So, function whatever we function we say is actually being performed by the cell what is present in that particular organ and that is why the cell is considered to be the structural as well as the functional unit of the cell. So, whatever the function you see from that particular organ could be performed by that particular cell as well which means a cell is the smallest unit which actually be able to perform all the functions. For example, in a human body we have different types of organs to perform the different types of tasks like you know heart is there for circulating the blood, liver is there for detoxification, kidney is there for the secretion of the byproducts, lung is there for you know inspirations, but the cell which is actually the structural and functional unit can be able to perform all these functions on its own because it has all the necessary infrastructure to perform all the functions.

Based on the cellular structures, cells are classified either as the prokaryotic cell or the eukaryotic cell. So, what you see here is a prokaryotic cell which is actually a bacterial cell and I have taken the two examples of the eukaryotic cell I have taken the example of the plant cell and I have taken an example of the animal cell. So, based on the structure and structure cellular structure the cells are classified into the prokaryotic and eukaryotic cell. In most of the cases prokaryotic cells are the single cell whereas the eukaryotes are

either single or the part of the multicellular tissue system. So, before getting into the detail of the structure of the prokaryotic or the eukaryotic cell list this is about the differences between the prokaryotic as well as the eukaryotic cell.

So, that you will understand what could be the differences what is going to happen and how the eukaryotic cell has evolved from the prokaryotic cell. So, what you see here is a table where I have listed the differences. So, this is the properties of the prokaryotic cell and this is the property of the eukaryotic cell size. So, first criteria is the size and the size is very small. So, prokaryotic cells are mostly in the range of the micrometer range whereas the eukaryotic cells could be of variable size they could be you know they could be up to the 40 micrometer in diameter.

So, they could be you know several sizes RBCs, microphages, kupffer cells and all those kind of things. So, they will be very different. As far as the genetic material is concerned the genetic material could be that genetic material is circular in the case of the cytosol and it is present as a free material which means it is not present in the bound form. Whereas the DNA in the form of a linear chromosome present in the well defined double stem membrane nucleus. So, no direct connection with the cytosol.

So, in the eukaryotes the DNA is present in the form of a chromosome and that is present in a well defined structure which is called as the nucleus and that nucleus is not directly under the contact with the cytosol. Then the replications as for the replication is concerned. So, replication means how you are actually going to make the another copy of your genome. So, the replication is done by the single origin of replication what is present in the case of prokaryotes. Whereas in the case of eukaryotes it is having the multiple origin of replications.

As far as the genes are concerned, so, the genes are the functional unit the genes are the functional part of the genome which are actually responsible for the production of different types of products or different types of proteins. So, for the gene there is no intron present whereas, in the case of eukaryotes you have a you introns are present in the eukaryotes. So, do not worry about this particular terminologies because these terminologies will be clear when we are going to discuss about the replication transcription and translations in a subsequent modules. Then the organelles there is no organelle the no membrane bound organelle is present in the prokaryotic system whereas, the membrane bound organelles are with the well defined functions are present. So, you have the different types of organelles you have the nucleus you have the mitochondria you have the chloroplast you have the endoplasmic reticulants and so on.

So, that all we are going to discuss in this particular module. Then the cell wall there is

a definite very complex cell wall what is present in the prokaryotic system whereas, in the case of eukaryotes except the fungi and the plant the eukaryotic cells are devoid of a thick cell wall which means the animal cells are devoid of the cell wall whereas, the fungi and the plants are going to have the cell wall. Then the ribosomes, so ribosomes are the protein machinery and they are actually going to be 70S. So, this is the kind of a you know parameter so that 70S and the whereas, in the case of the eukaryote it is the 80S. Then we have the transcription and the translation.

So, transcription and translations occurs simultaneously in the case of prokaryotic cell whereas, in the case of transcription in the nucleus. So, transcription is happening within the nucleus and the translation is happening within the cytosol as we said only already in the beginning that the nucleus is DNA the genome is present in a well defined nucleus and that is very far away from the cytosol that is why the transcription is and the translation is not happening in the same simultaneously transcription is happening in the nucleus and the translation is happening inside the cytosol. So, before so now let us start about the discussion about prokaryotic cells. So, the simple prokaryotic cell what is being shown here, this is a bacterial cell and the structure of the prokaryotic cell a simple prokaryotic cell is very simple and smaller than the eukaryotic cell as we said you know prokaryotic cells are in the range of micrometer whereas, the eukaryotic cells are very big compared to the prokaryotic cell. One of the classical difference between a prokaryotic cell and a eukaryotic cell that it has a well it has no membrane bound organelle including the nucleus.

So, you what you have is a cell where all the organelles are present within this cytosol and what you have here is the different types of organelle you have the flagella, you have the you know genomic DNA or the gene or the genome of the bacteria or the prokaryotic cell then you have the well defined cell wall, then you have the plasma membrane and then it has a protective capsule which is actually going to give the strength and protection and then we it has the ribosome which is called as the protein machinery. Then it also has the food granules and all other kinds of things and it has the pili then it has a cytoplasm and plasmids. So, let us discuss about all these the substructures what is present in the prokaryotic cell. So, the first substructure is the flagella. So, flagella is present in the those bacteria which are actually motile.

So, flagella is present in a bacteria and it is required for the motion within the bacteria. So, you can see that if a bacteria is present in a drop, it actually can use this flagella to swim around. So, flagella is attached to the bacterial capsule is a central feature of most of the prokaryotic cell especially the motile bacteria. It provides the motion or the locomotion to the bacteria and it is responsible for the chemotaxis of the bacteria. I am sure you probably are not aware of this terminology which is called as chemotaxis.

What is mean by the chemotaxis? Chemotaxis means the attraction of organism attraction towards chemicals. For example, if there is a sugar crystal, if there is a sugar crystal, then what the bacteria is going to see, it is actually going to move towards this sugar crystal because it is looking for that particular sugar crystal it wants to eat that. So, that motion that directed motion of a bacteria towards the particular chemical is known as the chemotaxis and how it will move, it is actually going to use this flagellum which is attached to the capsule. The movement of a bacteria towards a chemical gradient is known as the chemotaxis, which means once you have a sugar molecule here, it is actually going to be dissolved into the water and it is actually going to have a gradient. So because this bacteria can be able to have the senses, it can be able to sense this particular gradient and that is how it will move towards that particular food source.

It could be glucose, it could be any other molecule as well. So flagellum is a part of cell wall and motion is regulated by the motor protein present inside the cell. So flagellum is attached to the cell wall and inside it actually has the motor neurons like just like when in a humans we have the muscles, it is simply it has also be attached to the cellular machinery so that it actually can have the motor neuron motor proteins and that motor protein actually can change the flipping movement. So it that the flagellum is actually going to have the flipping movement, it is actually going to move like that. And the flagellar motion is an energy consuming process and it is governed by the ATPase present at the bottom of that particular shot.

It is made up of the protein which is called as flagellin and the reduction or the separation of the flagellar protein reduces the bacterial infectivity and ability to grow. So some of the bacteria also uses the flagellum even for accessing the different types of host and that is how they can be able to use this for reaching to the host. So they can also use for reaching to the host and that is how they can be infectious. So if you actually reduce the production of this flagellin protein and if somehow you compromise the flagellar movement, you are actually going to make the bacteria non-motile and that is how the bacteria is going to lose its ability to infect and that is how they will be not going to cause the disease. Now the second is the bacterial surface layer.

So as you can see the bacteria has a very, very complex surface layer because bacteria possesses the three anatomical barrier to protect the cells from the extreme external damages. So since the bacteria does not have the membrane bound organelles and it is very susceptible for the hypotonic lysis because bacteria is mostly been present in the water or hypotonic solutions. It has a very well defined anatomical barrier so that it can actually be able to withstand this. So what are those caps? What are the different layers you have? You have a bacterial capsule which is the outermost layer and it is made up of the high molecular weight polysaccharide. So what you see here is this is the bacterial

capsule and this capsule is required because it gives the protections and this is the outermost layer and then it is impermeable to the water or other aqua solvent and it is responsible for the antigenicity of the bacterial cell.

Then you have the cell wall which is present in the middle layer and the cell wall and it is responsible for giving its response to the gram staining and the third is the plasma membrane. So you have the three layers one is the capsule. So the outermost layer is capsule then you have the cell wall and then you have the plasma membrane and why the bacteria has such a complicated system because bacteria is always been present in a harsh environmental condition. It could be present in the water, it could be present in the strong acid solution, it could be present in a alkali solution, it could be present in a solution where a lot of chemical toxicants are present. So because these are things are there it actually is protecting themselves by using all these layers.

So capsule is a very, very thick layer which actually is not going to allow these chemicals to get inside the cell. Then we have the cell wall. So cell wall composition in gram negative and gram positive bacteria is different. The cell wall has different constituents and be responsible for their reactivity towards the gram stain. So we have the two different types of bacteria one is called as the gram positive bacteria and other one is called as the gram negative bacteria and both of these bacteria will have the different types of cellular cell wall compositions and because of that they will be differentially responsible for one of the classical stain which is called as the gram stain and because of the gram stain, they are being classified either as the gram positive which means the gram positive are actually going to give you the staining whereas they are going to be gram negative if they are not going to be give you the gram staining.

And based on this only the gram staining the bacterias are classified as a gram positive or gram negative. So let us see what are the different components are present in a cell wall. The outermost layer what you have is a peptidoglycan layer. So this is the peptidoglycan layer what you see here. So the peptidoglycan layer is very thick in the case of the gram positive bacteria.

So you see the multiple layers are present in the gram positive bacteria whereas it is very thin in the case of the gram negative bacteria and because of this only it is actually having a differential response towards the gram staining. peptidoglycan is a polymer of the NAG and the NAM. NAG is N-acetylglucosamine and the NAM is the N-acetylmuramic acid. So NAG and NAM are actually the sugar molecules which are present and these sugar molecules are connected to each other by a peptide by a beta 1, 4 linkage and the sugar polymers are attached because you see this it has you know sugar molecules like so these are the sugar molecules which are been attached to each other by a beta 1, 4 linkage and

alternatively you have you have the NAM block then it is connected to a NAG block and then it has a NAM block like that. So you have you see that it has a one layer then you have second layer you have third layer you have fourth layer and fifth layer and then these layers are been attached by the peptide chains which are composed of the amino acid L-alanine, D-glutamic acid, L-lysine, and the D-alanine which means it is actually a combination of the L and the D amino acid.

If you are not very aware very much aware of the L and D form then we are going to discuss that when we are going to talk about the amino acids. But these are the two different types of amino acids and you know that the L amino acids are more abundant in the nature compared to the D amino acids. So the peptide chain present in one layer cross linked to the next layer to form a meshwork which is responsible for the physical strength of the cell. So what you have is you have the NAM and NAG blocks and then the second layer is also having the same way and then these layers are actually been connected by the peptide chains and these that is why it is actually giving a tensile strain to the cell wall and that is why they are very, very, you know robust or they are very, very rigid in terms of accepting the outside molecules. And the peptidoglycan synthesis is targeted by the antibiotics such as penciline whereas the lysosine actually degrades the peptidoglycan layer by cleaving the glycosidic bond connecting the NAG and NAM to form the polymer.

So you have the two options if you want to destroy the cell wall what you have is you can actually have the antibiotics. So if you put the antibiotics what antibiotics is going to do is it is actually going to target the peptidoglycan synthesis. One of the classical example is the penciline. The other is option is that you can use an enzyme which is called as the lysosine and that lysosine is actually going to degrade the linkage between the NAM and the NAG. Lysosine is a very, very important enzyme what is present in our tears and as well as the saliva and that is how the tear and the saliva are actually protecting the humans from the bacterial infection because as soon as the bacteria goes irrespective of whether it is a gram positive bacteria or the gram negative bacteria the lysosomes are lysosine is actually cleaving the bond between the NAM and NAG and that is how they are actually destroying the cell wall and once they destroy the cell wall they are these bacterias are very, very susceptible for the osmotic damages.

So they will be very susceptible for the water and or they will be very susceptible for the tear like conditions and that is how they were actually going to get lysed and that is how they will die. So this is one of the strategies and that is how people are trying to develop many antibiotics which are actually going to work on the peptidoglycan synthesis. Apart from that the cell wall is also having the lipotychoic acid. So apart from the peptidoglycan layer you also have the lipotychoic acid which is present in the cell wall. So lipotychoic acid are only present in the gram positive bacterial cell wall and it is an important antigenic

determinants.

So aside this for lipotychoic acid our immune system is actually going to work and that is how it is actually going to produce a response. Then we have the lipopolysaccharides or the LPS. The lipopolysaccharides are only be present in the gram negative cell wall and it is important antigenic determinants. So compared to the lipotychoic acid which is only present in the gram positive bacteria in the gram negative bacteria you have the lipopolysaccharide and that lipopolysaccharide is a very very important antigenic determinants because that is actually going to induce the immune response in the humans. This is I have just given you a write up so that if you are interested you can actually be able to read about the gram staining.

So gram staining is a staining which is been you know developed by Danish scientist which is called as the Hans Christian gram and as I said you know gram staining is gram positive bacteria is taking up the gram stain whereas the gram negative bacteria are not taking up the particular stain. So if you want to be more interested about reading the gram staining you can be able to go through with this publication and as well as I have given you a small write up so that you can also go to this particular write up as well. So now let us move on to the beyond the cell wall. So apart from the cell wall you have they have the cytosol and the other organelles. So prokaryotic cells do not contain any membrane bound organelles.

The organelles are present in the cytosol such as the ribosome which is the 70S ribosomes and the genetic material whereas the electron transport chain and complexes are embedded within the plasma membrane. So within the plasma membrane you have the electron transport chain you will see the description about electron transport chain when we are going to talk about the mitochondria. Apart from that genomic material is present in the chromosome and as well as the extra chromosomal DNA. So prokaryotic cells contain the genetic material in the form of a circular DNA known as the bacterial chromosome. So that bacterial chromosome is different from the eukaryotic chromosome what is present in the eukaryotic cells.

It contains the genetic elements for the replications, transcription and the translations. Bacterial chromosomes follow a rolling circle model of the DNA replications. The genes present on the chromosome does not contain the non-coding region which is called as the introns and it is co-translated to the protein. Besides main circular DNA, bacteria also contain the extra chromosomal or extra circular DNA known as the plasmids.

So what you see here is actually a plasmid. These plasmids are called as the extra chromosomal DNA which means they are actually be important for the bacteria but they

are being present as a extra chromosomal DNA. Presence of plasmid containing resistant genes confers the resistance towards the known antibiotic. Exchange of extra chromosomal DNA between the different bacterial strain is one of the mechanism responsible for the spread of antibiotic resistance across the bacterial populations. So the plasmid is very important because the plasmid is the only genetic material which actually been exchanged between the different bacterial species and that is how they can be able to exchange their properties with the neighboring residues. For example, if you have bacteria has suppose 200 copies of a plasmid which is actually giving a conference the resistance against the antibiotic for example, antibiotic pencilene.

Then what it will do is it will actually going to give some of these plasmids to the another bacteria which is actually sensitive bacteria. So once it will be sensitive bacteria are actually going to receive these plasmids they will also going to be resistance for the antibiotics. That is why it is important that when people are working in the laboratories or when people are working in a biofirm industries or something these plasmids has to be the plasmid bearing bacteria which we people are generating while they are doing the recombinant DNA technology has to be destroyed very nicely so that the genetic pool of this plasmid should not go into the environment and that is how the if there will be an exchange of the genetic material or exchange of the plasmid between the two bacteria it is actually going to spread the antibiotic resistance even in the natural bacteria. And that is one of the mechanism through which the bacterias are acquiring resistance and they acquire the resistance very fast because the exchange of the plasmid material and that is why it is important to study about the plasmids. We have the so plasmid is a circular DNA and there are different forms of plasmids what is present when you are going to do the plasmids for the different types of treatment.

For example, if you take the so you have the bacterial plasmid is a double circular DNA molecule and it exists in the three different forms. If the both strands of the circular DNA are intact, it is called as a covalently circular DNA. So this is what you see here is a covalently circular DNA. Whereas if one of the strand has nick then it is acquired the confirmation of the open circular DNA. If you are actually going to put the nick in one of the strands like for example here, then it is actually going to acquire another confirmation which is called as the open circular DNA or the OC form.

This is called as the triple C form. This is called the OC form and the third is called as the super coiled form. During the isolation of the plasmid DNA from the bacteria covalently circular DNA loses few numbers of turn and as a result it acquired the super coiled confirmations. The interchange between these forms are possible under the in vitro or the in vivo conditions such as the DNA gyrase produces the additional terms into the circular DNA to adopt the circular confirmations. So, bacteria plasmid is actually

acquiring all these three confirmations under the in vitro or the in vivo conditions. And that is how they can be you know that side they different enzymes are working.

For example, if you take the circular DNA and if you put the DNA gyrase it is actually going to create the turns into this and that is how it will going to generate the super coiled DNA. But if you take the super coiled DNA and treated with the topoisomerase it is actually going to make you know reduce the turns and that is how the it is going to be turned into the closed covalently closed circular DNA. Let us see one of the plasmids. These are the bacterial plasmids which are very commercially been available or very been used in the recombinant DNA technologies. So plasmids are widely been used for the cloning of foreign DNA into the bacterial system as a host strain.

And this is the plasmid which is you know which has the different types of components. So one of the thing what you have here is the origin of applications. So origin of application is a place from where the bacteria is actually going to start its replications. Then it also has the antibiotic resistance genes. For example, here you have you see it has the antibiotic resistance genes which is for the ampicillin.

So if this plasmid will go to any bacteria it is actually going to give the resistance against the ampicillin the antibiotics. So the antibiotic resistance genes or the enzymatic gene is responsible for giving the phenotypic changes in the host after the entry of the plasmid. Apart from that you what you see here is an enzyme which is also being present within the plasmids. And because the plasmid has the origin of replication it has the antibiotic resistance and it has all these components they are independent and that is how they are very, very you know dangerous because they can be independent they do not be dependent on the nucleus for its replications or early activities. And that is why they can independently go to the new bacteria and the new bacteria is also going to have the additional features.

So, the first thing to remember is this plasmid is actually acquiring. Let us see how you we can be able to isolate the plasmid from the prokaryotic cell. So the plasmid isolation is a multiple step process it is having the many steps. So in the step one what you have to do is you have to collect the bacterial. So first you have to do is you have to take the bacteria you have to transform that bacteria with the plasmid or suppose the bacterial plasmid is present in the bacterial cell.

First thing is what you have to do is you have to grow these cells so that you have a sufficient number of bacteria. Then you will have to do what you have to do is you have to in the step one you have to do a centrifugation and then you have to resuspend these bacterial cell in a solution one. The solution one is actually containing the 50 millimolar

glucose, 25 millimolar tris SCL and 10 millimolar EDTA. So the method what we are discussing is called as the alkaline lysis method.

In the step two, you are going to do the alkaline lysis. So in the step two, you are going to do the alkaline lysis that alkaline lysis you are going to do with the help of the point two normal NaOH and 1% SDS and that is actually going to lyse the cells and it is going to denature the DNA. Ultimately, you are going to do the third step which is called as the renaturation. So in the renaturation is going to be performed by the potassium acetate solutions and the glacial acetic acid solution. And what will happen is that in this step, there will be the renaturation so that renaturation is actually going to renature the plasmid DNA but it will not renature the chromosomal DNA and because of that the chromosomal DNA is actually going to be discarded in the next step. So when you are going to do the centrifugation the chromosomal DNA since it is not been renatured it is going to be precipitate and it will go to be present in the form of pellet whereas the supernatant is going to contain the plasmid DNA.

Then in the step four, you are going to just do the purification of this plasmid. So you are going to do the deprotonations and that will result into the isolation of the plasmids. That deprotonation you are going to do with the help of the chemical which is called as the phenol chloroform isomer solutions and that is actually going to remove the protein so that you can be able to make the very highly quality high purity plasmid DNA. And in the step five you are going to resuspend that plasmid into an alcohol from the solutions. So that is what you are going to get in the step four you are actually going to have the plasmid as well as the protein and then what you are going to do is protein will be present in the precipitate that supernatant you are going to collect and then that supernatant is going to add the alcohol.

And once you add the alcohol the precipitate the plasmid is going to be precipitate and that is how you are going to isolate the pure plasmid and that pure plasmid can be used for the different types of applications like different types of application which we are not going to discuss. So let us give you a very real experience how you can be able to isolate the plasmids from the bacterial cell. So I will take you to my lab where my student is actually going to show you a very small demo and how you can be able to isolate and the plasmids from the bacterial cell. Hello everyone in this video we will show you how to isolate plasmid DNA using alkaline lysis method. For preparation of plasmid DNA we need resuspension buffer, lysis buffer and a neutralization buffer.

In addition to that we need isopropanol, RNase and ethanol. Resuspension buffer contains 25 millimolar tris and 10 millimolar EDTA and we have to add RNase at a final concentration of 100 microgram per ml. Lysis buffer contains 0.2 normal sodium

hydroxide and 1 percentage SDS.

Neutralization buffer contains 3 molar potassium acetate, pH 6.0. For isolation of plasmid DNA we need at least overmet grown culture with OD of 3.0. So this is already cultured one. We have to harvest the cells by centrifugation. In these files we have to centrifuge 11000 rpm for at least 1 minute to get the cells precipitated.

Now we got the cell pellet. We can proceed for alkaline lysis method to isolate plasmid DNA. In first step we are going to add resuspension buffer which contains RNase. Mix thoroughly until all the cells suspended in resuspension solution. After the cells got suspended completely now we have to lyse the cells using strong alkaline condition that is 0.

2 normal sodium hydroxide and also 1 percentage sodium dodecyl sulfate. Now we have to gently flip the tube in order to lyse the cells completely. We can keep in this condition for up to 5 minutes but not more than 5 minutes which will degrade the plasmid DNA and also genomic DNA will come out and it will interfere with the mini-grid. In next step we have to neutralize the sodium hydroxide using neutralization buffer to prevent any further degradation. After adding neutralization buffer you can see there is a white precipitate that means all the proteins precipitated by neutralization buffer. You can flip the tube 2-3 times completely precipitate all the remaining proteins.

Now the solution part contains our plasmid DNA and all the precipitated one contains genomic DNA and also the proteins from bacteria. Now we have to transfer the white clear supernatant to another tube. This clear supernatant contains plasmid DNA. Now we have to precipitate this plasmid DNA with the isopropanol followed by washing with the 70 percent ether.

We can see white precipitate in the solution. Now we have to centrifuge it, collect the white precipitate and wash with the 70 percent ether. After precipitating plasmid DNA with the isopropanol we will get a pellet of plasmid DNA. Now we have to wash that pellet.

We wash this pellet with the 70 percent ether. Now we got the pellet. We have to air dry the pellet and dissolve it in DIL dries of water or TE buffer. We will keep leave at room temperature till the etheral gut evaporated. Next we will have to easy the process of manual alkaline lysis method. There are several kits available from commercial vendors. The basic difference between alkaline lysis method and the kit basement method is kit basement method contains silica basins columns where lysis lysate which containing plasmid DNA binds through these beads and after washing whatever the unwanted

components are there they will elute out and we will elute the plasmid DNA in TE buffer or water.

The composition of the lysis buffer is same as previous method and also neutralization buffer, every buffer contains same composition. But in commercial kits we have one extra wash buffer which will remove any unwanted contamination and give pure DNA. So when you see this demo what you could see is that the students was discussing about the all the 4 or 5 steps what we have just discussed and after these steps what you are going to get you are actually going to get the plasmid like this. So what you see here is the 3 forms of the plasmid you have the covalently closed circular DNA which means that the TLC forms you are going to have the OC forms and then you also going to have the super coiled form. So what you see here is this is actually the closed circular covalently closed circular DNA.

This is the open circular DNA and this is the super coiled DNA and whereas since we have not used the RNAs you are also going to see the some amount of RNA because RNA is also present in the bacterial cell. So what we have discussed so far we have discussed about the bacterial prokaryotic cells we have different we have discussed about the structure of the prokaryotic cell. We have discussed about how the cell wall is thick in the case of the gram positive bacteria whereas it is thin in the case of the gram negative bacteria apart from the cell wall you also have the capsule and as well as the plasma membrane and all these barriers anatomical barriers are making the bacterial cell very very resistance for the environmental changes or the chemical what are present in the environments. So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss about the eukaryotic cells. Thank you. .

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati

Module - 01
Basics of Biological system (Part 1)
Lecture - 02
Cellular Structure (Eukaryotic cells)

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati. And in this module, we are discussing about the different types of cells. So, cell could be of two based on the structures, the cell could be of two different types, it could be a prokaryotic cell or the eukaryotic cell. In the previous lectures, we have discussed about the prokaryotic cells, we have discussed about the structure of the prokaryotic cell and how the different components are present in the prokaryotic cell. We have discussed about the cell wall, we have discussed about the different types of analytical barriers what are present in the prokaryotic cells.

And then we also discuss about the chromosomal DNA and as well as the plasmids. And in the previous lecture, we have also discussed about how you can be able to isolate the plasmids from the bacterial cell and how you can be able to use them for the different types of applications and how the plasmid is actually giving the, you know, spreading the resistance or the different types of properties. So, in today's lecture, we are going to discuss about the eukaryotic cells. So, when you talk about the eukaryotic cell, we have discussed we said you know, we are going to discuss about the two types of eukaryotic cells, one is called as the plant cell, the other one is called as the animal cell.

So, let us start our discussion about the eukaryotic cells. So, the higher eukaryotes have the multiple organs to perform the specific functions whereas, the you know, the whereas, the cell is the structural as well as the functional unit of the life and it contains all the necessary infrastructure to perform the all the functions. So, besides we have the different types of cells like we have the eukaryotic cell, we have the eukaryotic cells and within the eukaryotic cell you have the plant cell and the animal cells. So, before getting into the detail of what are the different organelles and what are the structure of the eukaryotic cell, it is important to understand the difference between the plant cell as well as the animal cell. So, the structure of the eukaryotic cell, the eukaryotic cells are much more complex and it contains many membrane bound organelles to perform the specific functions.

It contains a nucleus isolated from the cytosol and enclosed in a well defined plasma membrane. So, one of the classical feature is that the eukaryotic cells are actually containing the membrane bound organelles to perform the specific function and then you have the two different types of eukaryotic cells, you have the plant cell and you have the animal cell. So, let us see what are the difference between the animal and the plant cells. So,

these are the properties of the plant cell and these are the properties of the animal cells. So, what you see here is the cell wall, the cell wall is present in the plant cell and cell wall plant has a very robust and good cell wall which is made up of mostly the cellulose whereas, the cell wall is completely absent in the animal cell.

Then the size the plant cells are very large compared to the animal cell and the animal cells are comparatively small. Then the chlorophylls, the chlorophyll is the light pigment, chlorophyll is the light pigment what is present in the plant and it is completely absent in the animal cell. There are exceptions where the animal cells are also having the chlorophyll like for example, the Euglena. So, Euglena is animal cell but it also contains the chlorophyll. Then we have the vacuoles.

So, there are large vacuoles which are present in the plant cells and these the purpose of these vacuole are that they want to collect the food material or sometime they also collect the toxic substances because the plants are much more exposed to the toxic substances whereas, in the case of the animal cell it contains the small and the small number in many number in the animal cell. Then we have the mitochondria. The mitochondria are few in the case of plant cell because the plant cell are not very motile the plant cells are static like you have seen that the trees are present in the soil. So, they do not move around that is why they do not require a large quantity of energy for motion and that is why they do not have the very huge number of mitochondrial cells. Whereas in the animal cell you have the huge amount of animal cell because animal cell are very motile.

So, because of the animal cell they are motile they are they need more amount of energy and that is why they have the more number of mitochondria. Then the lysosomes, the lysosomes are almost absent in the plant cell whereas, they are present in the animal cell. Then we have the glycosomes, the glycosomes are present in the plant cells, but they are absent in the animal cell. Then you have the cytokinesis, cytokinesis means the cell division this is we are going to discuss in due course. So, the cytokinesis is the cell division, cell division is always by the plate method in the case of plant cells, whereas it is going to be by constriction in the animal cell.

So, these are the some of the notable differences between the animal cell as well as the plant cell. But overall the basic property of these 2 cells are very, very same actually they both have the plasma membrane they both have the many of the organelles what are present. So, let us start discussing about the different types of organelles what is present in the eukaryotic cell. And we are going to then take up the some of the organelles what are exclusively been present in the plant cells and so on. So, let us start our discussion with the first organelle and that organelle is called as the cytosol.

Cytosol is nothing but the water or the liquid part what is present inside the cell. So, you see the cell which is there and it has a nucleus in the center and in between whatever you see is actually called as liquid part whatever you see is called as the cytosol. So, cytosol is

the liquid part filled inside the cell and it contains the water, salt, macromolecules which means it contains the proteins, lipids, RNA. It has an array of microtubules fiber running throughout the cytosol to give the vesicular structure to its destinations. So, within cytosol you have the different types of microtubules running.

So, these microtubules are actually making the road within the this cytosol and on these roads only the vesicles are moving and that is how they are actually delivering their content just like as you are ordering the material from the Amazon or flip cards and then the guys are coming to your place. But how they are coming they are coming because there is a road there is a road to your home and that is how they come using these roads. So, these roads are made up of the microtubules which we are going to discuss in our subsequent lecture and they will use these roads and that is how these vesicles you can imagine these vesicles are the courier guys which are coming to your home and they are actually delivering to your place and how they are delivering that is a part of the very, very extensive vesicular structure system and they are actually having a very, very well defined vesicular transport system and vesicular transport system is actually going to help the cells to deliver the material to their destinations. Beside this the cytosol exhibits the salt to gel transitions and such transitions regulates the multiple biochemical and cellular processes. So, salt and gel transition is actually making the either the cytosol more thick or more thin and because of that, the localized salt and gel transition is actually going to make the substrates more concentrated or less concentrated and because of that they can be able to enhance or they can be able to change the rate of these reactions very nicely.

So, apart from these functions the cytosol has many of these functions. So, cytosol has no well defined functions, it serves as a medium to exchange the material between the different organelles. It play in you know various processes such as it is actually we appraises where the signal conduction or that translation is going to take place. So, cytosol has the cellular machinery which is responsible for the translation or translation means the production of the proteins. So production of protein is called as the translation and that is happening within the cytosol.

Now, one of the things which people are very oftenly use or interchangeably they use the term which is called as the cytosol or the cytoplasm and there is a big difference between the cytosol and the cytoplasm and it is very important that you should get a clarification. So, cytoplasm is actually the liquid part. This is the liquid part whereas, the cytoplasm is actually consists of the cytoplasm and all other organelles. So, if you take the cell, it has a nucleus, outside this nucleus whatever you have that is called as the cytoplasm that contains the cytoplasm and it also contains the different types of organelles. So, that is a very, very important to you understand that you should say cytosol when you are talking about the liquid part and you say about the cytoplasm when you are talking about the content what is outside the nucleus.

So, let us move on to the next organelles and the next organelle is known as the nucleus.

So, nucleus is the central processing unit of the cell and it is homologous to the processor in a typical computer. Nucleus is also called as the master of the cell because it actually regulates it actually say in give the instructions to the cell what it should do actually. So, big why it is so because the nucleus is containing the genetic material and that genetic material has all the information it and that genetic material is actually going to say what the cell has to do actually. So, nucleus is a very, very well present inside double membrane structures.

The liquid which is filled inside the nucleus is called as a nucleoplasm and the nucleoplasm is actually contains the different types of molecules like it is a VCS liquid which contains the nucleotides and the enzyme to perform the replication and the transcriptions and as well as the DNA damage repair system. So, the nucleus contains the nucleotides which is required for the replication as well as for the transcription and then it has the replication machinery. Replication machinery means it is actually be going to make the DNA another copy of DNA. So, that will required when the cell is going through with the cell division, then it contains the genetic material the DNA in a complex fashion involving the several protein which are called as the histone proteins. So, compared to the bacterial cell, the DNA is present in a complex with the proteins and that these proteins are called as the histone to pack into a nuclear bodies or to the chromosome.

So, the chromosome when we say about the chromosome it is a authentic chromosome what is present in the eukaryotic cell. The chromatin in the eukaryotic nucleus, so chromatin like the DNA content is present in the two forms either it can be euchromatin or the heterochromatin. So you can have the chromosome which is going to be either be present as the euchromatin or the heterochromatin. So, euchromatin is a part of the chromatin where the DNA is loosely packed and it is transcriptionally active to form the messenger RNA. So, euchromatin is actually active whereas the heterochromatin is non active or inactive which means the eukaryotic cell has the mechanism so that they can be able to make some portion of their genome active and some portion of their genome inactive.

Why it is so because they do not want to expose their whole DNA for the cellular machinery and as well as for the different types of toxicants and because of that they can be able to protect. So, the DNA what is present in the heterochromatin is actually going to be more densely packed and because of that it is getting protected from the any kind of damages. So, they will be protected from the damages and it is transcriptionally inactive. So the portion what you are going you what you require for example, if there will be a requirement of the insulin molecule for example, if we require insulin, then that insulin gene is actually going to be present in the euchromatin and that is how the insulin is going to be transcribed within the nucleus and that is how the insulin production is going to start in the cytosol. Since the nucleus is present inside this particular double membrane structure, it is actually well protected from the cytosol or well separated from the cytosol by the very, very complex structure and that complex structure is called as the nuclear membrane.

So, the nucleus in eukaryotes are present in a double layer of membrane known as the nuclear envelope. So, what you see here is this is the nuclear envelope where you have the nuclear pore and this is the only portion through which the molecules can come in or the synthesis synthesized molecule can go out. The outer membrane of the nuclear envelope is continuous with the rough endoplasmic reticulum and has the ribosome attached to it. The space between the two membrane is called as the perinuclear space and the nuclear envelope is often has the nuclear pore and as per the calculation and average nucleus has the 3000 to 4000 pore per nuclear envelope. So, you have these type of nuclear pore and the nuclear pore is a 100 nanometer diameter and it contains of different types several types of protein.

It is a gateway for the transfer of material between the nucleus as well as the cytosol. So, you cannot just get anything into the nucleus because then that will in that that is the case there are many you know DNA damaging agents what are present in the cytosol. So, they will not get into because the nuclear pore is actually going to you know have a mechanism so that they can be able to discriminate between who will be allowed to get in and who will be allowed to go out of this DNA. The RNA formed after the transcription from the DNA within the nucleus and it will move out from the nucleus into the cytosol to the nuclear pore. Similarly, the protein from the cytosol crosses the nuclear pore to start or initiate the replications, transcription and other processes because you require the signal from the cell also you require the cellular machinery from the cell with you so that you can be able to perform the different types of functions for example, the replication and transcriptions and that is also been governed or that is also been completely been controlled by the nuclear pore.

Now, let us move on to the next organelle and the next organelle is called as the mitochondria and the mitochondria is also called as the powerhouse of the cell because it is responsible for energy production. So and that is why we if you recall when we were talking about the plant versus the animal, the animal is having the more number of mitochondria because they require more energy compared to the plant cell. And as the organelle is actively involved in the generation of a molecule which is called as the ATP to run the cellular activities. Mitochondria is a double layer membrane bound is organelle with a different structural properties outer membrane is smooth. So, this is what you see here this is a double membrane structure and the outer membrane is smooth.

So, this is the outer membrane what you see here is a smooth membrane and it covers the complete organelle with a large number of integral protein which are known as the porin. So, on this outer membrane you have the proteins which are called as the porins and the porin allows the free movement of molecules less than 5000 Dalton within and outside the mitochondria whereas the large molecules are or the protein moves into the mitochondria through the transporter involving the signal peptides known as the mitochondrial targeting sequences. So, it is not like just like the nucleus it is not like that anything will be allowed

to get in or anything will be allowed to go out. There is a small size like if it is a size of 5000 Dalton that is the molecule which is going to be allowed getting through the porins. But if it is bigger than this suppose you have the 50000 Dalton then that protein has to be go through with a well defined mechanism which means it should have a signal peptide so that and that signal peptide should be of a mitochondrial targeting sequence.

So, do not worry about these targeting sequence and all these kinds of terminologies because that we are going to discuss. Inner membrane is folded into a membrane projection which are called as the cristae. So, what you see here is the inner membrane which has been folded in the form of the cristae and the cristae occupies major portion of the membrane and house machinery for the anaerobic oxidation as well as the electron transport chain to produce the ATP. Due to the presence of the inner and the outer membrane the mitochondria can be divided into the two compartment first in between the inner and the outer membrane. So, this is the compartment and known as the intermembrane space and the second inside the inner membrane known as the matrix.

So, this is the matrix what you see here this is the matrix. The protein present in the intermembrane space have a role in executing the programmed cell death or the apoptosis which we are going to discuss in a subsequent modules. Matrix is the liquid part present in the innermost of the mitochondria and it contains ribosomes, RNA, DNA enzymes to run the crepe cycles and the other protein. So, this protein what this portion what you see here is called as the matrix and the matrix is actually going to have the or the cellular machine it contains the ribosome so that you can have the protein production because you know that the nucleus has the mitochondria has its own mitochondrial DNA. So, it can be able to do the transcription and translation and that is how it can be able to produce the proteins.

Then it also has the RNA, DNA and it also has the enzyme for the running the crepe cycles and all of the proteins. Mitochondrial DNA is a circular DNA it has a full machinery to synthesize its own RNA which means it can have the ability to synthesize messenger RNA, ribosomal RNA and as well as the transfer RNA and the proteins. A number of differences exist between the mitochondrial DNA and difference and the DNA what is present in the nucleus. One of the classical difference is that the mitochondrial DNA does not contain the intron. So, it does not contain the intron compared to the DNA what is present in the nucleus.

Now, let us talk about the cellular the ATP generation machinery. So, this is the ATP generation machinery which is called as the which is present in the micro which is present in the plasma membrane and the system what produces the energy is called as the electron transport chain which is actually a very, very huge complex and that is why these are labeled as the complex 1 to complex 5 and these are all complexes are the integral proteins and they are present in the inner membrane of the mitochondria. During the metabolic reactions such as glycolysis, crepe cycle produces large amount of reducing equivalent in the form of the NADH₂ and as well as the FADH₂. Electron transport chain processes the

reducing equivalent and the flow of electron through the different complexes. So, what happened is that the electrons are moving from the complex 1 to the complex 4 and that causes the generation of the proton gradient across the membrane.

So, what happened is it is actually going to accumulate the large quantity of the proton then the proton expelled in the intermembranes space return back to the matrix to the complex 5 and this is what you see here the complex 5 which is called as the ATP synthase and what happened is that the big when the electrons are moving from the complex 1 to complex 4 they are throwing the proton from the this side to this side and that is how there will be accumulation of the proton and then when this proton comes back to the matrix part matrix side it is utilizing the this complex or this ATP synthase and because of that the ATP synthase is coupling the ADP plus PI to generate the ATP and that is because ADP synthase is a very very big complex it has the F1 unit and as well as the FO unit and it is actually. ADP synthase is a mushroom shaped multimeric protein complex mainly composed of two proteins FO and F1. FO is a membrane bound portion so this is the FO which is a membrane bound portion and F1 is the complex present into the lumen towards the matrix. FO F1 complex of mitochondria harvest protein motive forces to catalyze the phosphorylation reaction ADP in responsible for generating the ATP. Now what are the different reactions what are the different function of the mitochondria so the first and the most important function of the mitochondria is that it is involved in the ATP productions then the second is it is actually be responsible for the generation of the active species to kill the infectious organism.

So when there is the entry of the infectious organism within the mitochondria or within the cell it actually induces the production of the reactant species. Oxygen species means any molecule which actually contains the single electron so because of this single electron they are very very toxic. So because they are toxic they are actually being used as a weapon to kill the microorganisms. Then the mitochondria is used to track the tray of a family you know that the mitochondria if you see the production of or the how the sexual reproduction occurs the sexual during the sexual reproduction the you have the two species you have the sperm and you have the ovum. Ovum is actually a complete cell it contains all other organelles so it contains the mitochondria also whereas the sperm is only containing all other organelles but only contains the DNA.

So when there is a fusion of the sperm and ovum it takes the DNA part from the sperm whereas it contains all other organelles from the ovum. Because of that when the zygote is being developed it actually contains the mitochondria and that mitochondria is from the mother side that mitochondria remain constant because even if this particular individual will again go for the sexual reproduction that mitochondria is actually going to carry forward and that is why you can be able to use the mitochondria to track the family tree. Then the mitochondria is also having a role in the programmed cell death or the apoptosis. So apoptosis is what we are going to discuss in the subsequent modules but I have given a very short write up so that it will actually help you to understand what is apoptosis. So

apoptosis is a programmed cell death involving the series of events involving cellular module metalloprotease known as caspases in an adverse event of exposure of the cell to the cytotoxic agent or to the environmental conditions.

It activates the cell surface signaling to activate the cytosolic proteases. In addition it disturbs the mitochondrial membrane potential to cause the release of a protein which is called as the cytochrome c. As you remember when we are talking about the transport of material within the cytosol and outside we said that the pore in is only allowing the 5000 Dalton proteins. But if there will be a membrane potential problem if there will be a drop in the potential then that will allow the leakage of the proteinaceous substances from the mitochondria and one of such protein is called as a cytochrome c. And once the cytochrome c is going to be released from the mitochondria it is actually going to activate the different types of DNA is like and that is how the DNA is are actually going to destroy the DNA what is present inside the nucleus and ultimately it is going to destroy that particular cell.

And this is actually a very programmed death and that is why this is called as the apoptosis. Now let us understand the first process how the mitochondria is actually synthesizing the ATP. So understand and to explain this process the people have post forward a theory which is called as the chemo-ismotic hypothesis. So what it says is that the F1 particle so you remember it has 2 particle one is FO particle the other one is F1 particle. FO is the integral membrane protein which is only responsible for holding the F1 and it also gives the path for the protons to enter whereas the F1 is actually a multimeric protein complex and it also it contains the 3 different proteins.

It contains a protein which is called alpha subunit it contains the beta subunit and it contains a gamma subunit. So this is the gamma subunit what is present in the center and the function of the gamma subunit is that it actually only transfers the complexes from the alpha to beta. So what you see here is that these alpha and beta proteins are or the subunits could be present in 2 different conformations either it could be present in the loose binding conformations or the tight binding conformations. So when they are present in the tight bind conformations they will bind that particular molecule and very tightly. So what will happen in the beginning in the beginning the ADB and the PI is actually going to come and bind to the the loose binding side.

So once they will bind to the loose binding side the gamma subunit is actually going to take these molecules and put it into the another subunit and once they will reach to another subunit the ADB and PI are actually going to mixed with each other and that is how they are actually going to form the ATP. As soon as it forms the ATP it will again been flipped to another subunit and that subunit is open complexes and that is how it is actually going to you know it is going to release the ADB and then it is actually going to take up the new molecules of ADB and PI. So this this cycling event is going to continue for many rounds and that is how it is actually going to keep giving you the ATP synthesis. So I have requested one of the professor in MRC and based on my request I think he has shared a movie which

actually explains these processes very nicely. So Professor Volker is we should be very grateful to him that he has shared this very good animated schemes.

So in this scheme they have shown how the ATP synthase is actually synthesizing the ATP utilizing the different types of complexes. So what you see here in this movie that this cyan color portion is actually the FO particle and all these what you see is F1 particle. This is what you see is the gamma chain whereas this blue yellow is the alpha chain and this is the beta chain. So what when you will see that when the it is actually utilizing the proton motive force it is actually rotating and when it is rotating it is actually accepting the ADB and PI which is a in the loose binding site and once the loose binding site is rotating it is converting into a tight binding site and that is how the ADB and PI mixed together and it generate the ATP. Again there is a rotation and because of that the it is getting converted into the loose binding site and that is how the ADB which is formed in the previous cycle is going to be released.

So this is a very very good movie to understand the whole process and that is how you will be able to appreciate how these events are actually been happening. So what you see here is actually it is showing how the ADB and PI are combining with each other. So you see these are different bonds through which the ADB is combining with the PI and that is how the ADB is formed and that is how it is actually been released from the system. So this is all about the ATP synthesis and its mechanism by the ATP synthesis and what we have discussed so far we have discussed about the eukaryotic cell and within the eukaryotic cell we have discussed about the differences of the eukaryotic cell with the plant cells and then we have discussed about the nucleus we have discussed about the cytosol and we have also discussed about the mitochondria. Now in our subsequent lecture we are going to discuss about the some more organelles from the eukaryotic cell and then we also going to discuss some more interesting aspects related to the cell.

So with this I would like to conclude my lecture here. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 01
Basics of Biological system (Part 1)
Lecture-03 Cellular Structure (Eukaryotic cells)

Hello everyone, this is Dr. Vishal Trivedi from the department of bioscience and biotechnology IIT Guwahati and what we were discussing we were discussing about the cells and what we have discussed is that the cell is the structure and the functional unit of the life. In the previous two lectures we have discussed about the prokaryotic cell and as well as the eukaryotic cell in the within the prokaryotic cell we discuss about the different parts of the bacterial cells we talk about the flagella we discuss about the genetic material of the antibacterial cell and we also have discussed about the cell wall and we also discuss very briefly about the gram staining. And then the in the previous lecture we discuss about the eukaryotic cells and what we have taken is we have taken the two eukaryotic cells the plant cell as well as the animal cells and we discuss about the several types of differences between the plant cell as well as the animal cell. And subsequent to that we discuss about the cytosol, we discuss about the nucleus, we discuss about the mitochondria. So, in the previous lecture we started our discussion about the animal cell and within the animal cell we discuss about three organelles we discuss about the cytosol, we discuss about the nucleus, and then we also discuss about the mitochondria.

In each nuclei, in each of these organelles, we discuss about their mechanism, their role in the cell cellular physiology. And then we also discuss about the different types of structural details of that particular organelles and so on. So now in today's lecture, we are going to start discussing about the some more organelles from the eukaryotic cell. So let us start today's lecture.

So the first organelle what we are going to start discussing about the chloroplast. So chloroplast is present in the plant cell and it is completely absent in the animal cell except that it is also present in euglena which is considered to be a primitive animal cell because the euglena has the two abilities to synthesize the food and it also can have the ability to trap its prey and it also can be able to take the food from the external sources. So chloroplast which are found in the plant algae and the other lower invertible animals such as euglena. Contrasting to the mitochondria, chloroplast has the outer membrane and inner membrane and then the light pigment containing innermost thalakoid membrane. So what you see here is this is a typical structure of the chloroplast.

What you see here is it has the outer membrane, then it has the inner membrane and inside the this inner membrane, you have this thalakoid membrane. So these thalakoid membranes are actually containing the light pigments. The outer membrane is porous to the small molecule but the protein or the large molecules are transported by the TOC. The TOC is a stand for the translocosin, locon on the outer chloroplast membrane

complexes. So similar to the mitochondria, if you remember the mitochondria is having the porins right which are allowing the proteins or even small molecules which are of lower than small 5000 or 1000 compared to here also the outer membrane is porous to the small molecule but it is not porous to the large protein or to the and that although large protein has to be moved through the facilitated transfer.

The movement of material of passed through the outer membrane enters into the inner membrane through the another complex which is called as the TIC which is called as the translocon onto the inner chloroplast membrane. So this is the TOC which is present onto the translocon which is present onto the outer chloroplast membrane whereas the TIC which is present on is a translocon which is the present onto the inner chloroplast membrane. In between the outer and the inner membrane is the intermembranous space which is filled with the liquid and liquid the inner membrane of the chloroplast further folds into the flattened membrane system such as the thalakoid. So this is just a simple this is the structure wise the chloroplast is following the similar kind of structure what is being present in the mitochondria. If you recall the inner membrane in the mitochondria is getting folded and it is forming the cristae and all other kind of structures whereas in the case of chloroplast it is actually forming the thalakoid membrane.

So these thalakoid membranes are actually containing the photosynthetic pigments and these pigments are actually being responsible for harvesting the sun energy and that is how that energy can be utilized for the dark reactions. So let us understand about the photosynthetic pigments. So the photosynthetic machinery such as the light absorbing pigments and the electron carriers and the ATP synthesizing machinery is present onto the inner membrane as a integral protein complexes. So these are the thalakoid membrane where you have the integral membrane proteins and all these integral membrane proteins are actually having the complexes which are responsible for the light absorbing complexes like the PS1 as well as the PS2 and then it also has the electron carriers like the cytochrome c and other kinds of electron carriers like the q electron carriers and then it has the ATP synthesizing machinery similar to the mitochondria. So it also has the ATP synthase which also going to participate into the ATP productions.

Thalakoid membranes are arranged into a stack of coins which are called as the granum. So all these thalakoid membranes so this inner layer is actually being folded into the thalakoid membrane and then these thalakoid membranes are stacked to each other. They will be stacked one after other and that is how they are actually going to form a coin like structure and that coin like structure is called as the granum. These granums are actually going to contain these light absorbing pigments, electron carriers and the ATP synthesizing machinery. The granum throughout the chloroplast are connected by a tubule to share the material.

So these granums are present in a chloroplast. So you see that if you have this like a chloroplast so the inner membrane is actually going to fold to form one set of granum then it is going to be connected and then it is going to form the another set of granum and then it is going to be connected and it will form the third layer of granum. So these molecules are going to be connected to the tubules and that is how they are actually be

able to share the material between the different granums. The overall structure of the chloroplast is similar to the mitochondria but it has few significant structural and biochemical differences. For example, the thalakoid membrane contains the photosynthetic green color pigment which is called as the chlorophylls.

So let us discuss some more about the function and the major function of the chloroplast is that it is actually going to be participate into the reaction which are called as the photosynthesis. What the photosynthesis mean is that photosynthesis is a very simple complex structure. Photosynthesis means you have to do the synthesis and you are going to utilize the energy from the sun which you are going to insulate the light energy which means you are going to use the light energy to synthesize so that is called as the photosynthesis. So let us see how the photosynthesis is happening. So photosynthesis is a simulation reaction involving the carbon dioxide and water to produce the sugar in the presence of solar energy or the photons to catalyze the fusion reactions.

So this is the typical photosynthesis reactions where you have the carbon dioxide, water and then you also require the energy from the sun and that is actually going to be combined to give you the sugar and it also going to give you the oxygen. So that sugar can be utilized for the plants for its own growth and as well as this sugar can be stored in the form of fruits which actually we are going to which are which other animals or other organisms are going to consume. The photosystem present onto the thylakoid membrane consists of the two photosystems. You have the photosystem 1 which is called as the PS1 and then you also have the photosystem 2 which is called as the PS2. Now these two photosystems are working in accordance with each other so that the electrons or the light energy what they are going to absorb from the sun is actually going to be utilized for the generation of the ATP.

So the purpose of the photosynthesis is that it is actually going to be used for synthesis of the two molecules. It is going to be utilized to synthesize the ATP and it is also going to be utilized for the reducing equivalent which is called as the NADPH. This ATP is going to be utilized is going to be synthesized by the molecule which is called as the ATP synthase whereas the NADPH is actually going to be formed by the electron transport chain or the electron transport system. And both of these molecules the generation of both of these molecule it requires the energy and that energy it is going to get from the sunlight. So PS2 is actually going to absorb the photon from the solar energy to excite the electron to the higher energy state and catalyze the water breakdown into the proton and oxygen.

So this is what it is going to happen. So the first complex what is going to respond to the sunlight is actually the PS2 and PS2 is actually going to take up the sunlight and that is how it is actually going to catalyze the breakdown of the water which is called as the water lysis and that it is going to generate the proton as well as the oxygen. This electron pass through with the and it also going to have the electron. So the electron which are going to be produced during this water lysis is actually going to be passed through from the multiple electron carriers and during this electrons are exported out of the thylakoid membrane into the lumen. So this is what you see here is it is going to do

the photolysis of the water that is actually going to generate the proton as well as the oxygen.

And on the other hand the electrons which are going to be excited from the PS2 are actually going to be carried forward throughout the lumens and throughout the thylakite and that is how it is also going to be utilized. The proton pass through the ATP synthase and the return back into the stroma to generate the ATP. So what happen is that the proton are actually going to be accumulated onto the this side and then they will actually going to pass through to the ATP synthase and that is how it is actually going to generate the ATP into the lumen. The electrons from the PS2 is eventually been received by the PS1 and been excited under the absorbing proton from the sunlight to high energy state. So that is why the proton the electrons are going to be pass through the different electron carriers and then it will reach to the PS1 and at the PS1 they will reach to the PS1 and again the PS1 is actually going to again receive the sunlight and then that is how it is actually going to excite these electrons and ultimately these electrons are going to be utilized for production of the NADPH.

The energy associated with these electrons are used to generate the NADPH into the stroma. So in the stroma what you are going to generate you are going to generate the two molecules you are going to generate the NADPH and you are also going to generate the ATP. So ATP and the NADPH both are actually been utilized for the dark reactions. Hence as a result of photosynthesis the solar energy is been trapped by the photosynthesis apparatus to generate the two molecule one is ATP which is the energy currency and the NADPH which is called as the reducing equivalent into the lumen. Both of them are being used to run the Kelvin cycle to assimilate the environmental carbon dioxide to form the sugar.

Now this is what is going to happen. So these two molecules are going to be utilized. So the purpose of the ATP photosynthesis is that it actually want to synthesize the ATP and the NADPH and then these two molecules are going to be supplied into the stroma where they have the enzymes for the Kelvin cycle and that is how the Kelvin cycle is going to run into the C4 plants and that is how it is actually going to synthesize the sugar molecule. So ultimately the carbon dioxide is going to be fixed by the plants into the sugar molecules and it is actually going to oxygen. That oxygen is going to be used by the animals for respiration and this sugar is actually going to be used by the plant for its own growth and the extra sugar is going to be stored in the form of fruits and that also is going to be consumed by the other animals.

Now let us move on to the next organelles. So the next organelle is the organelles of the vesicular trafficking. You know that every cell just like as we are also having a very very good trafficking system so that you know that what is the destination of this particular road. If you want to go from for example if you want to go from Guwahati to Mumbai you know that what will be the roads I should take to reach the Mumbai or suppose I want to go to Delhi or Kolkata or any other place. Similarly if you want to distribute the material within the cell then also it also has the vesicular trafficking system.

So there are organelles which are responsible for distributing the material within the cell. This could be for the plant cell or it could be for the animal cell. This material either could be the food particles or it could be the signaling molecules. So this could be anything. So the main function of these organelle which are actually the organelle which are part of the vesicular trafficking is to manage the distribution of the material whether it is a food particle or the protein which are a part of the signal transduction throughout the cell.

There are three different organelles such as the endoplasmic reticulum, Golgi apparatus and the lysosomes which coordinate work together to maintain the vesicular transport of the material across the cell. Eukaryotic cell take up the solid material from the outside the cell through a process which is called as the endocytosis. So if it is taking up the solid material then it is called as the endocytosis whereas the uptake of the liquid is known as the pinocytosis which means during the nutrition during when the cell is taking up the nutrition it can actually take the particulate matter that process is called as the endocytosis. If it taking the liquid for example if it taking the water or any other kinds of vitamins and minerals and all those kind of molecule then it will be called as the pinocytosis. Similarly the material is secreted out of the cell which is called as the exocytosis.

So inside entry is called as the endocytosis if the cell is producing some byproducts which are not good for the cell then it is also going to throw the cells and that process is called as the exocytosis. In addition the intravascular system delivers the protein synthesized in the endoplasmic reticulum to the different organelles. Apart from these two these three processes like two processes where the cell are actually going to receive the material if it is solid then it is called as the endocytosis if it is liquid then it is called as the pinocytosis and if it is a byproduct then it is called as the exocytosis. Apart from these three movement of the material distribution of the materials you can also have the distribution of the material to the different organelles. For example you know that the all the proteins are being synthesized either inside the endoplasmic reticulum or inside the cytosol.

But these proteins probably may not be required for that particular organelle it may be required for the lysosome it may be required for the mitochondria it may be required for the chloroplasts. So that movement is also be a responsibility of these organelles which are part of the vesicular trafficking. During the endocytosis the material present outside the cell binds to the cell surface receptor and trapped it in a membranous structure which is called as the endosome. The endosomal vesicles are fused with the lysosomes to form the endosome. In late endosome with the help of the lysosomal enzyme material is digested and then the endosome is fused with the Golgi bodies and deliver the content for the further distributions.

In the similar manner during secretions the vesicular originate from the Golgi bodies and fused with the plasma membrane to release the content. So this is what you see here. Here we have shown the all the three processes one is the endocytosis. So if it is a food particle it is going to be take up inside it will be going to engulf and then it is actually

going to be first present into the early endosomes. These early endosomes when they will fuse with the lysosome but it is going to be present in the cytosol.

Then it is actually going to form the late endosome and then these late endosomes are actually going to be fused with the Golgi complexes and then the Golgi is going to process this particular material what is being taken up from the outside and that is how it is actually going to be delivered to the raffinoplasmic reticulum or it is actually going to be given to the other organelles. Same is true for the if suppose the something has to be secreted out like for example if something is has to be exocytosis or something has to be secreted out then that material is going to be come out in the form of the vesicles and then these vesicles were eventually going to fuse to the plasma membrane and then this material is going to be go out. Same is true for the exocytosis where the Golgi is going to pack this material in the vesicles and then these vesicles are going to fuse with the plasma membrane and then it is actually going to release this particular content. Now let us study these organelles individually and understand their functions. The first organelle which is be a part of the vesicular trafficking is the endoplasmic reticulum.

Endoplasmic reticulum is nothing but the roads which are present inside the cell. So what you see here is the endoplasmic reticulum is present just outside the nucleus and it is forming a road like structure. It is forming a road throughout the cell. So if you want to send a material which is suppose for the mitochondria then these roads are actually going to go to the mitochondria and that is how it is actually can deliver that material to the mitochondria. So the vesicular network starting from the nuclear membrane and spread throughout the cytosol constitutes the endoplasmic reticulum.

There are two different types of endoplasmic reticulum which are present in the cell. You have the rough endoplasmic reticulum which is actually having the protein machinery attached to it which is ribosomes. So you have the rough endoplasmic reticulum and the smooth endoplasmic reticulum. Rough endoplasmic reticulum is having the ribosomes which are attached to it. So because of this ribosomes they are their appearances look like as a rough endoplasmic or rough surfaces.

So the rough endoplasmic reticulum has ribosome attached to it and it gives the rough appearance whereas the smooth endoplasmic reticulum is devoid of the ribosomes. Protein synthesis on the ribosome attached to RER is sorted into three different categories such as integral membrane proteins, proteins for the secretions and the protein designated for the other organelles. So the protein what is been synthesized inside the endoplasmic reticulum actually falls under the three different categories. Number one is the protein which is a part of the integral membrane proteins. Number two the protein which is for the secretions and the number three the protein which is for the different organelles.

Protein are synthesized with the N signal peptide and these signal peptides are recognized by the signal recognition particle on their target organelles. So the protein which are designated for the different organelles are synthesized with a signal peptide. So the signal peptide is nothing but kind of a address. So they are actually having a address. So you can imagine that if I want to post a letter from here IIT Guwahati to IIT

Madras then what I will do is I will take the letter I will have a letter and then I will write
an address on this.

Similarly if I have a vesicle and if I want to send this vesicle to the mitochondria what I will do is I will put the mitochondrial localization sequence. And that is what is called as the signal peptide. You remember that when we were talking about the last time when we were talking about the mitochondria that the pore in will not allow the entry of any proteins which is beyond the 500 Dalton. But if the protein is of beyond the 500 Dalton then that protein has to have a mitochondrial localization sequence. So you can put a tag like if you put a signal then this vesicle will go to that destinations.

For example if a protein is synthesized with a signal peptide for the mitochondria it will attach to the signal recursion particle and the receptor onto the outer membrane to deliver the protein. The proteins without any signal peptide tags remain in the cytosol. So any protein which does not have any kind of tag is actually going to remain within the cytosol. Now what will be the function of the endoplasmic reticulum? So the first function is that it is involved in the synthesis of the steroid hormone within the gonad cells.

Then it is required for the detoxification. Remember that the endoplasmic reticulum is a part of the vesicular transmittance. So it actually can do the exocytosis and that is how it is actually going to participate into the detoxifications. Then it also can do calcium sequestrations and that is how it actually can have the calcium signaling. So if it gives the calcium signaling the endoplasmic reticulum is actually going to release the calcium into the cytosol and that is how it is actually going to start the calcium signaling. Then it is also important for the synthesis of the protein phospholipids and the carbohydrates.

It is possible for the protein sorting for the different organelles and it is also responsible for the protein modification such as glycosylation. So some of these things we are very, very complicated and we are not going to discuss. For example, the glycosylation itself is a big topic so that we are not covering in detail in this particular course. Then we talk about the next organelle and the next organelle is the Golgi bodies. The Golgi bodies are actually being forced to visualize by a metallic stain which is called as the Golgi stain invented by the Camellio Golgi and it is made up of the, so Golgi is made up of the flattened disc like cisterny arranged in a stacked manner to give three distinct zones.

So this is what you see. This is the Golgi bodies where you have the disc like structure. So disc like structure which are attached to each other and that is how it is going to have a Golgi bodies. You have the three different zone within the Golgi bodies. You have the cis zone, you have the medial zone and then you have the trans zone.

So this is the starting point. You have the cis zone. So cis phase is actually receiving the material or vesicle from the endoplasmic reticulum. So this side with the side from which it actually receives the material from the ER is called as the cis phase or the cis cisterny. Whereas the middle portion is called as the medial Golgi where in medial Golgi you are actually going to have the all the processing. It is actually going to have the

covalent modification with the sugar. So it is going to do the different types of glycosylations and all those kind of modifications.

And then the top portion what you see here is actually the trans Golgi. The trans Golgi is actually is the phase of the Golgi towards the plasma membrane and this site is actually going to release the sorted vesicles whether these vesicles are going to be for different organelles or whether this is for the plasma membrane which means whether these vesicles are for the secretory pathway or whether these vesicles for the other mitochondria and for their designated organelle or to the plasma membrane. So these are the functions of the Golgi bodies. You have the protein sorting. So produce in the medial Golgi the proteins so it will actually receive the protein what has been synthesized by the endoplasmic reticulum then that protein are going to be sorted within these medial Golgi and then by sorting these proteins are actually going to be modified by differentially.

They are going to be tagged with the different types of destinations. For example, it can be a mitochondrial localization sequence, it could be chloroplast localization sequence, it could be some other kind of localization sequences even for the Golgi itself if Golgi want to get some protein it also has to put a Golgi localization sequences and ER localization sequences. Although this protein is coming from the ER but it cannot remain in within the ER it has to be received from the Golgi bodies. All the material will go into the Golgi then it will be going to be sorted out and then it is going to be tagged with the particular address and then subsequently it is actually going to be delivered to that particular all male. For example, you know if you are in a home if you are in your home and if you send a envelope or if you send a letter what happened this letter first go to the GPO.

Then from GPO it is actually will go to the different you know postal address or this will go to the postal office. From the postal office it will go to the postman and then postman is actually going to deliver it to the destination. So same is true for the vesicular trafficking. If you are in the home this is actually the ER. So where the synthesis is happening then what will happen this is your letter.

So this is a protein. So this is a protein. Now this protein will first come to the GPO. GPO is nothing but the Golgi bodies. Now from the Golgi bodies it is actually going to be sorted. It will actually going to be sorted as per the destination.

For example, there are parcels which will go to the Mumbai. There are parcels which are going to the Delhi. There are parcels which will go to the you know Kolkata or other cities. So at this point it is actually going to be sorted and it actually going to have some kind of stamp that okay this will go to the Kolkata. This will go to the Mumbai.

This will go to the Delhi and something like that. So then it will reach to the Delhi office and then from Delhi office it will be given to the postman and then postman will actually going to give you the destinations. So this is the very important this is a Golgi bodies. So even if the letter has to be come back again it has to be come back again to

the your home for example. Then also the GPU has to tag accordingly then only it will come back. For example, if you are sending a letter to your neighbor it will not go directly from your place to that neighbor.

It will go through this process. It will go to that particular postal office and then it will come back to your that neighbor house. Same is so that is the function of the this vesicular organelle what is involved in the vesicular trafficking. Apart from that the Golgi is also involved in the proteolysis so where it is also going to degrade the proteins. Now we are talking about the third organelle which will also be responsible or be a part of the vesicular trafficking and that organelle is called as the lysosomes.

Lysosome is an organelle which is discovered by the D.D.O.V. and they are popularly known as the societal back because the lysosome is filled with the different types of hydrolytic enzymes and its inner liquid is very acidic. So due to their role in the autophagy, autophagy means eating yourself. So autophagy is means eating yourself which means if you say you might have seen many people who are chewing their nails that is autophagy actually that is that you are chewing your own body. So same is true for the cell also when the cell does not produce the enough energy because it has to you know it is not getting a nutrition from outside then what will start doing is for example if suppose it has the 300 copies of some organelle.

So what it will do is it will start utilizing the 100 copies. So the it will actually going to be work with the 200 or copies of that particular organelle and the 100 copies it will going to you know destroy and that material it is actually going to use for its nutrition. That process is called as the autophagy but this is a societal pathway right and that is why the lysosomes are known as the societal backs. Autophagy is a cellular process probably operate in cells during starvation to meet their energy requirements. Lysosomal lumen is extremely acidic and contains the proteases cytosolic cytolytic enzymes to degrade the ingested material. So if you have a lysosome and if you give any any molecule whether it is a protein whether it is a DNA whether it is the bacteria viruses any kind of molecule it is actually going to degrade and it will going to generate the proteins or peptides.

That is why the lysosome has very well defined function it will degrade it will degrade the ingested food material for delivery throughout the vesicular systems. So if you know if you take the food particles from outside it is going to be delivered to the lysosomes and that lysosomes is doing to what is going to do it is going to degrade that food particles so that it would be present in the form of the simple molecules and that simple molecules it is going to deliver. It is also been present in the defense cells and it is going to work as a defense organelle also so it is actually going to destroy the pathogenic bacteria viruses and yeast fungi and all kind of pathogenic bacteria and then it is also going to degrade the old protein. So the major part of the lysosome is that it is actually going to recycle the material it is going to recycle the cellular materials and as well as it is going to recycle the outside material. So if there is a bacteria if it goes into the cell it is actually going to destroy that particular cell so that bacteria will be given to the into the you know to the lysosomes by the very well defined process and that anyway we are

going to discuss when we are going to discuss about the cellular processes.

And we are going to talk about phagocytosis that time we are going to discuss in detail how the bacteria or viruses or all these infectious organisms are going to be delivered to the lysosomes and then how the lysosomes are actually degrading these bacteria. Now let us move on to the next organelle and the next organelle is called as the plasma membrane. So plasma membrane is nothing but the external membrane and a plasma membrane is made up of the two molecules it is made up of the lipids and it is made up of the proteins. So you know that the protein the lipid has a head and then it also has the aliphatic chains hydrophobic chain and these head molecules which are called as is actually the polar and the these chains are hydrophobic and because of this particular type of amphipathic character all these heads are actually arranging themselves and the lipids are the chains are actually arranged inside. So if you say this so if you put this under the aqueous environment it is actually going to form a membrane like this and that is how the plasma membrane is made up of the lipids as well as the proteins.

So what you see here is this ball like structures are actually the polar head groups and what you see here is this tail like structure is the hydrophobic tails and they will be arranged and sandwiched so they will make a sandwich like structures and within these sandwich structures you are going to see the different types of proteins. These proteins could be the integral proteins or they could be peripheral proteins. So what you see here is this is the integral protein because this is present throughout the plasma membrane whereas this is what you see is actually a integral protein which is either be present onto the outer surface or to the inner surface. Apart from that the plasma membrane is also going to have the different types of receptors like for example this is a receptor this is a receptor it can also have the channels like it can have the transporter as well as the channels. These transporters and channels are actually going to use for delivery of the food particles or delivery of the small molecules.

Apart from that it also could have the different types of the molecules which are being attached onto the lipids membrane and that is actually going to a part of the antigenic role. So they are going to be the antigenic molecules which are going to be attached. So these are sugar molecule which are being present on some of these peripheral proteins and that is responsible for giving the antigenic features to this particular plasma membrane. Now what is the function of the plasma membrane? So function of the plasma membrane is that it is actually going to protect the self from the external infections.

Then number two is it is actually be responsible for entry or exit of the molecules. So it is a part of the regulatory system so that it actually will allow the entry and exit of the molecule because the plasma membrane are semi-permeable. So they will allow some molecule semi-permeable so they will be selectively permeable. So they will be having some mechanism so that they will be very selective whether they will be allow some molecule to enter or not that is actually going to be decided by the plasma membrane. Apart from that the plasma membrane is going to have the different types of receptors so they will be actually going to use that for many purpose. The receptors could be taking

up the for taking up the food for example you can have the receptors which is for the taking of the food for example you can have the LDL receptor so that LDL receptor is going to take up the LDL which is a lipid actually from outside and that LDL receptor is going to take up the LDL and that will be utilized for the cell for its nutrition.

Similarly you can have the insulin receptor. So insulin receptor is not going to use for detecting the insulin what is present in the blood and that is how it is actually going to lower down the blood glucose. Number 3 the receptors are also going to be a part of the defense mechanism. So some of the receptors are going to function as you know the recognition particles or sometimes they are also going to work as a defense mechanism. So they will be going to sense the external molecules and they are actually going to derive the responses from the cell accordingly.

So and apart from that the plasma membrane also has the transporters. These transporters are actually going to be used for the different types of delivery of the molecules or the delivery of the water or the solutes and small molecule as well as the big molecules. So overall the function of the plasma membrane is that it wants to regulate the material entry and exit from the cell. So with this we have discussed about the eukaryotic cell what we have discussed we have discussed about the differences between the two different types of eukaryotic cell we discussed about the plant cell and as well as the animal cell and we also have discussed about the different organelles what are present in the eukaryotic cell. Initially we have discussed about the cytosol then we discuss about the nucleus and we also discuss about the mitochondria and in this particular lecture we have discussed about the chloroplast then the organelles of the vesicular trafficking system we discuss about the endoplasmic reticulum we discuss about the Golgi bodies and we also discuss about the lysosomes and at the end we have also discussed about the plasma membrane and its functions. So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss about the cell cycle and how the cell is dividing and increasing its number and some more aspects related to the cells. So with this I would like to conclude my lecture here. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 01
Basics of Biological system (Part 1)
Lecture-04 Cellular Fractionation (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT, Guwahati. And what we were discussing, we were discussing about the cellular structures of the prokaryotic cell and the eukaryotic cell. And in the previous two, three lectures, we have focused more on to the structure and as well as function of the different parts, what is present in these type of cells. So when we were discussing about the prokaryotic cell, we have discussed about the plasma membrane, we have discussed about the cell membrane, we have discussed about the plasmids and we have also discussed how you can be able to isolate the plasmids and how the plasmids can actually be able to exist in three different forms. And subsequent to that, we have also discussed about the eukaryotic cell.

When we were talking about the eukaryotic cell, we have discussed about the plant cell and as well as the animal cell, we have discussed about the differences between the plant cell and the animal cell. And we have also discussed about the different types of membrane-bound organelles what are present in the eukaryotic cell. So we discuss about the nucleus, we discuss about the mitochondria, chloroplasts, glycosomes, Golgi bodies, endoplasmic reticulum. And at the end, we have also discussed about the plasma membranes.

And while we were discussing about these cell organelles, we were also discussing about their structures, functions, and what is their contribution in regulating the different events within the cell. So in today's lecture, we are going to discuss about how you can be able to study these individual organelles. So although the membrane-bound organelles are not present within the hyperkaryotic system, we can still be able to isolate the different fractions from the prokaryotic cell and we can be able to study those fractions. So let's start today's lecture. So before getting into the detail of these aspects where we are actually going to study how you can be able to fractionate the cells, how you can be able to isolate a particular fraction, let's understand that we have the prokaryotic cell, we have the eukaryotic cell.

And within the eukaryotic cell, you can have the unicellular eukaryotic cell such as yeast or you can have the multicellular prokaryotic eukaryotic cells such as animal cell, plant cell or the fungi. All these cells are very different from each other. And that's why

the first question comes if you want to isolate a particular fraction of these cells, how you can be able to grow these cells. And if you want to grow these cells, you are actually going to provide them the necessarily nutrition. So let's see what are the different nutrients, what is available for these cells.

So when we talk about the growth medium, so growth medium of a host organism is required to produce the four major basic biomolecules. One is protein, second is carbohydrate, third is lipid and the fourth is the DNA or the RNA. Now when we want to synthesize the proteins in organisms, the proteins are made up of the amino acids and these amino acids are actually going to be made up of the five different types of atoms such as carbon, hydrogen, oxygen, nitrogen and sulfur. In comparison to that, the carbohydrates, carbohydrates are commonly known as sugar. They are made up of the three different types of atoms, carbon, hydrogen and oxygen.

Then lipids, lipids are made up of the fatty acids. So lipids is a derivative of the fatty acids and the fatty acids are made up of the different types of atoms such as carbon, hydrogen, oxygen, phosphorus and sulfur. DNA and RNA mostly are made up of the nucleotides and these nucleotides are made up of the carbon, hydrogen, oxygen, nitrogen and phosphorus. The role of these biomolecules are different. For example, the protein is a building block.

So it actually requires for making the different types of structures. I am sure you might have seen the different types of proteins what are present in the different parts of the prokaryotic as well as the eukaryotic cell. Remember that the different types of receptors what are present onto the plasma membrane or different types of protein complexes are present in the mitochondria for ATP synthesis and the kinds of things. So those are the building blocks. Whereas the carbohydrates and lipids are mostly being used as a source to provide the energy.

Energy in terms of the burning of these molecules so that you can be able to produce the ATP and then this ATP would be you know the going to be the molecule which going to be supplied the or which can actually going to carry the high energy bonds and that actually going to carry the energy from one part to the other part. Lipids are also been mainly been used for the energy. So if suppose our organism want to move from one part of the dish to the another part of the dish or one place to another place, it actually requires a nutrition energy for the locomotion and the energy comes from the mainly from the carbohydrates and the lipids. Similarly we have we require the DNA and RNA. So for DNA, RNA is always been a part of the genomic content of the organisms either it can be DNA or RNA.

So anyway we are going to discuss in detail about the genomes and that time it will be clear in what are the organisms where you are actually going to have the RNA as the organisms. So apart from these three molecules so actually you require a nutrient source which actually can be able to provide you the carbon, hydrogen, oxygen, nitrogen and sulfur and as well as the phosphorus. So if you take a nutrient media or if you require a growth media which can actually be able to supply you the carbon, hydrogen, oxygen, nitrogen and sulfur and if that is actually going to you know so that is good enough for a cell to assimilate the different types of building blocks such as amino acids, sugar, fatty acids or nucleotides and ultimately they are actually be able to generate the biomolecules such as protein, carbohydrates, lipids and DNA and RNA and at the end they are actually going to use these molecules for running their metabolism, running their different types of activities and that's how they are actually going to use this for their growth. Apart from these two these major molecules you also require the minor quantity of the minerals or the vitamins. So considering these you can actually be able to design the different types of growth medium which is going to be used by the host cell or by the prokaryotic or the eukaryotic cell for their growth as well as the deployment.

So these are the different types of atoms what is required to be provided by the growth medium and considering the requirement of these atoms you can actually get these from the different types of biomolecules. So what are the different types of constituents which can be a part of the microbiology medium and mediums. So you can actually have the amino nitrogen which is actually going to be so amino nitrogen is going to be provided by the pectins protein hydrolysates in views and extracts then you require then you can actually have the growth factors. So growth factor can be provided by the blood, serum, e-sect or vitamin or NAD then you require the energy source that can be provided by the sugar, alcohol and carbohydrates and you also require the buffer because maintaining, you know, the pH of the media is also very important for the growth. So that can be done by the phosphate, acetate and citrates.

Then we can also require the mineral salts and metals. So like minerals, metals and cofactors. So that can be provided by these kind of metals. Then you also require the selective agents such as the chemicals or antimicrobials or dye. So that can be more relevant when you want to use these microorganisms for other kinds of molecular biology manipulations.

Then you require the indicator dye. So in some of the media, you also require the indicator dye so that you can be able to see whether there is a change in pH or not. So for that you can use the phenol red or neutral red. And then sometimes you also require the solid media. So solid media you require the agar or gelatin or silica gel.

So considering all these constituents or the source, you can be able to make the different types of microbiology media. So you can have the M9 minimal media, you can have the M63 minimal media, you can have the Luria Bertani and LB media, you can have the LB Linux media, you can have SOB media, YT media, Terephi broth media, Super broth and T pi g. And these are the composition what I have given and these are the bacterial species what you can actually use for cultivation. So you can actually use the M9 media for cultivation of the E.coli and other kind of E.

coli strains. Similarly you can use these kind of media for making the competent cells and all that. So all these terms are probably will be new to you but it will actually go and you will understand these when we are going to discuss about the different types of molecular biology aspects. And for example, this YT media, YT media is required to the production of phage productions and all that. So how the first question comes how you are actually going to prepare a microbiology media. So for preparing a microbiology media, what you require is you require constituents.

So for example, if I take an example, you can actually require the peptone, you can require the E-shearad, you require the NaCl and so on. So for preparation of microbiology media, you dissolve the component in 1 litre of distilled water and then you cover the top of the flask with a cotton plug and autoclave the solution at 121 degree Celsius for 20 minutes. So that you are going to do in the autoclave. And what are the precautions while you are making the microbiology media. So ratio of media volume and culture flask that is very important for aeration.

So for example, if I am going to make the 250 ml medium, then it has to be prepared in 1 litre flask. So that you can actually have the enough aeration because when you are going to grow the bacteria, bacteria require the oxygen and you cannot open the cotton plugs. So bacteria require a certain amount of oxygen for gaseous exchange so that it can take the oxygen and it can actually be able to release the carbon dioxide and that is actually going to decide what will be the final growth of these particular bacteria. Then you require that the media components should be hygroscopic and while weighing avoid the moisture. So store in a cool and dry place.

So you can understand that if there will be a moisture then the amount of actual powder is going to be lower because you are actually getting the weight of the water as well. So that is actually going to make the media less nutritious for the bacteria. Then while autoclaving, open the autoclave only after it is cold because the water is boiling inside the autoclave and it is actually going to have the vapors. So you have to take those kind of precautions. Then you are actually going to avoid the charring of the media components.

So charring means you are actually going to, you should not burn the media component because if they got burnt, then for example, if the glucose got burnt, right. So glucose is actually going to be converted into carbon. And this form of carbon is not suitable for any kind of biological system to assimilate, right. So that you cannot use. Then you also require the solid media should be poured in a plate once it is cold because if you pour the media which is very hot, then it is actually going to destroy the other kinds of additives.

For example, if you are adding the antibiotics or if you are adding the selection pressures and all that, then it is actually going to destroy or inactivate them. So the various antibiotic or nutrient supply should be added to the media when the temperature is less than 50 degrees Celsius. So this is all about a critical aspect, how you can be able to prepare the microbiology media. We have prepared a small demo clips to explain you how you can be able to prepare the microbiology media in the lab and how you can be able to do the autoclaving and all that. Hello everyone, in this video we are demonstrating how to prepare bacterial culture prep.

For preparing culture prep, we need three components. One is peptone, yeast extract and sodium chloride. For 100 ml of culture prep, we need 1 gram of peptone, 0.5 grams of yeast extract and 1 gram of sodium chloride. For 100 ml of peptone, we need 1 gram of sodium chloride.

After weighing, we have to clean the spatula and keep it in original position. And during weighing, care should be taken to avoid contact with any of these media components. After weighing the media components, we have to dissolve them in double distilled water. So initially we are dissolving in 80 ml of distilled water. Once the components completely dissolved, we have to make up the volume up to 100 ml.

While it is till, we have to prepare cotton plugs for the flask. For preparing cotton plugs, you have to take one thick layer of sheet of cotton. If you have two hands, fold like this. Once the media dissolution is complete, we have to pour into the flask.

We have to pour up to 100 ml. So we use only one third of the place. Remaining space is empty. This is used for aeration purpose and also ensure proper autoclaving. In order to check whether the components are autoclaved or not, the media is autoclaved or not, we use stability indicator.

This is paper based stability indicator. We have to paste on to the flask and we have to autoclave. If the autoclave is properly finished, then we will see the white strips turning into the black one. So this is the indication of the autoclave flask. Now the media

components are completely dissolved.

Now we have to pour into the flask. Once the media preparation is complete, we have to sterilize the media in order to use further applications. This is the typical autoclave where you can see temperature and pressure indicator and these are the pressure knobs and this one is quick pressure release knob. You can use when you are in a hurry, you have to use this one, but I will prefer not to use this one. Let it go on its way.

We have to turn on the autoclave. So you can see here the bulb is glowing. Before keeping the media components to autoclave, make sure that the heater inside the autoclave submerged with water. Now I am going to keep the media components in the basket which we use for the autoclaving. Then keep this one inside the autoclave. While closing the autoclave, make sure that you are closing in opposite direction.

Once the pressure and temperature reaches 121 degree Celsius and 15 kPa pressure, you have to hold on that point for 20 minutes. Then you have to turn off the machine, let it cool down and remove the components. The same procedure you have to while opening, you have to open in opposite direction. To conclude the video demonstration, we have discussed how to prepare bacterial culture media and how to prepare cotton bags and autoclave it. During culture weighing of the media, we have to make sure that the media component should not be exposed to air because those substances absorb the moisture and become liquefied.

So another thing is that for cotton block preparation, we have to take a single layer of cotton, then we have to fold it. After autoclaving, we should not release pressure in a single shot. Let it go and come to normal pressure, then we have to open autoclave. So with this, the video is over and thanks for watching. Now when we talk about the eukaryotic cells, right because you also require the studying the different fractions of the eukaryotic cell, the simplest eukaryotic cell is the yeast.

So you can actually be able to use some of the yeast media. So you can actually be able to use the CSM media, you can use the YPD broth, you can use the YPGAL, you can use the standard minimal media, and you can use the yeast nitrogen based media and all that. And I have given the composition for one meter, what you can actually be able to use. And this is the application. So for example, the YPGAL media is the standard media for S-survey C for omitting the glucose repurposition and all that.

Similarly, YPD broth media is commonly been used for the yeast media for the maintenance and propagation of pichia pectores and S-survey C, right. So pichia pectores and S-survey C are the most common yeast which are being used in the

laboratories. And method of preparation is almost the same as per the media composition, the consequences are being added in the 900 ml of water and autoclave, then you allow the media to cool down and then you add the 50 ml of filter sterile 40% glucose so that the final concentration will become 2% and adjust the final volume to a 1 liter if necessary. So this is about the yeast media. Now, moving to the next media, which is the mammalian cell culture media, right.

So initially, we discussed about the basic media, right, which is going to be utilized by the unicellular eukaryotic cell. Now you can use the more specialized media which is going to be used for the more specialized mammalian cells, we can still, so for the mammalian cell culture media, which is you see this, this is a DMEM media, right. So this is the DMEM media, which is prepared and the composition is that you are going so DMEM constituents are actually very, very high. So you require large amount of amino acids, different types of sugar, different types of other kinds of so it when you are going to buy the DMEM media from the vendor, it is actually going to provide you the powder.

So you require the 13.4 grams of the powder for 1 liter, then you require the sodium bicarbonate which is 3.7 grams per liter. And then you also require the fetal bovine serum which is FBS so 10%, normally you're going to get the FBS of 10, 100%. And then you also require the antibiotics.

So you are actually going to use the 1%. So 100 from the 100x you are going to use a 1x. How you're going to prepare the cell culture media? To explain the method of media preparation, we are taking the example of DMEM media. So add 13.4 grams of dry powder media into the water and mix it to dissolve it completely.

Then you add the 3.7 grams of sodium bicarbonate, mix completely and adjust the pH to 6.9 to 7.1 using the one normal NAH or one normal HCL. Finally, you add the cell culture grade water to the media to bring it to the final volume, sterilize the media using a sterilized filter, merminate filter with a pore size of 0.

22 micron. Then you are actually going to add the supplements such as the antibiotic and serum can be added to the sterilized solution using the aseptic techniques within the biseptic cabinets. This is all about the theoretical explanations. We have also prepared a small demo clips to explain you how you can be able to prepare the media within the laboratory or within the biseptic cabinets. Hello everyone. My name is Bhaumadh Rafi, a research student at IIT, Guwahati in biosciences and bioengineering department.

In this video, we are going to demonstrate how to prepare cell culture media for mammalian cells. For preparing cell culture media, there is a step by step process. First

we need to weigh the components of the media and dissolve it in required amount of water. Then we need to set the pH using ES clip and then we need to filter the media using 0.

22 micron filter to make it aseptic. For further use, we can also adequate the media and store it in 4 degrees. In this video, we will be demonstrating how to prepare mammalian cell culture media. For that purpose, we need aerosol media which is DMEM, double cross modified, eagles media as to the high glucose. We need FBS, fetal bovine serum and we need antibiotic cocktail comprised of streptomycin and anti-cellulone. The vessel media provides inorganic materials and amino acids which are preferred for basic development of cell and the FBS is used for providing both factors to the cell.

We cannot autoclave this media because it might degrade the components of the media. For that purpose, we use 0.22 micron filters. This is a 250 ml bottle top filter.

After we have packed the filter, we have to keep it for autoclaving. We have to keep it for autoclaving. After adding media, we need to stir it on a magnetic stirrer for the components to dissolve completely. We can either use double distilled water or milli-q water but double distilled water is more preferable as it contains more ions than milli-q water. After the media components have dissolved completely, we need to set the pH of the media.

For that purpose, either we can use pH meter or pH strips. In this case, we cannot use pH meter as the bulk of the pH is sensitive to the media components and may get corroded. After the media components have dissolved completely, we will be able to set the pH of the media. We need to adjust the pH of the media. The bright red colour indicates that the pH of the media is in the range of 7.

2 to 7.4. If the colour of the media turns purple, then it indicates that the media is acidic. If the colour of the media turns yellow, then it indicates that the media has become acidic.

Now, we will be checking whether the media falls in the range of 7.2 to 7.4. After the media has been set, we now need to filter the media inside the biosafety cabinet as we have added the constituents in their non-aseptic condition. After the media components have been completely dissolved and the pH has been set, we now need to sterilize the media using membrane filter media. For that purpose, we use class 2 biosafety cabinets which are used for handling mammalian cell cultures. So, this is a typical biosafety cabinet in which we perform the filtration for media.

This is the control panel which is used to operate this machine. This is the on and off

switch. This is the switch for normal light. This is the switch for UV light. Now, we are going to filter the media. For that purpose, we need a suction pump which can be connected to the bottle top filter.

This suction pump is for the purpose of extracting the air from the bottle top filter so that it can be used for the purpose of extracting the air from the bottle top filter. After the media has been filtered, we now need to add a PS and antibody in order to make it 10 percent FPS containing serum. So, we have seen how to prepare cell culture media for mammalian cells. We have seen how to prepare the cell culture media for mammalian cells. So, while you are going to prepare the media, you also have to consider the different types of different types of precautions.

For example, the pH of the media. So, pH of the media should not be between 7 to 7.4. This is the physiological pH. It should not be above to this or it should not be below to this because that is actually going to adversely affect the growth and as well as the other kinds of features of the cells. Filtration should be performed at a very low speed so that you should not feel while you are filtering, you should not compromise the pore size of the membrane.

And if you are going to filter at a very high speed, you are actually going to generate the shear stress. And because of that shear stress, you are actually going to increase the pore size. You know that the pore size of these membranes are 0.22 micron. So if you are actually going to bring very high speed, the pore size may grow up and it becomes like 0.

25 micrometer or it could be even 0.4 micrometer depending on how fast you are growing. So if you become 0.4 micrometer, some of the bacteria which are lower to this value are actually going to enter into your media. Apart from that when you are spinning or when you are filtering at a very high speed, you are also going to take up the air and air is nothing but the which contains the microbiologies. It can contain the pathogenic bacteria.

And you are actually going to add the serum which is heated activated serum. Then you also require the antibiotics and you are actually going to check the contaminations. So when you prepare a media, what you are going to do is you are going to take a small dish and you are actually going to put the 10 ml media and put it into 37 degree Celsius for 2 days. And if that is nothing will grow, right. So then you can observe this plate and if this plate does not show any kind of bacterial growth, then you can imagine that your whatever the media you have prepared is perfect and it can actually be able to good to use for the mammalian cell culture. Apart from that you can also use the insect cell

culture media and this is the recipe what I have provided.

The media preparation method and everything is almost the same as what we have discussed. So this is all about the cell biology and how we can be able to use the different types of techniques for the cell fractionations. So what we have discussed so far, we have discussed about the cell propagations, we have discussed about how you can be able to prepare the different types of media. So we discussed about the microbiology media, we discussed about the yeast media, we discussed about the mammalian cell culture media and we have also discussed about the insect cell media. We have also shown you the couple of demo so that you can be able to prepare these media in your laboratory as well such as the osmotic lysis, thermolysis, sonifications, mechanical methods and so on. And once you have broken the cells, you are actually going to get the different types of fractions whether it is a prokaryotic cell or the eukaryotic cell.

You can be able to use differential centrifugation and as well as the density gradient centrifugation to separate the different types of organelles and then you can utilize these fractions for your subsequent studies. For example, you can actually be able to study the mitochondria, you can be able to isolate the nucleus and do the different types of experiments. So with this, I would like to conclude my lecture here. In our subsequent lecture, we are going to discuss some more aspects related to the biological system. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 01
Basics of Biological system (Part 1)
Lecture-05 Cellular Fractionation (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT, Guwahati. And what we were discussing, we were discussing about the cellular structures of the prokaryotic cell and the eukaryotic cell. And in the previous two, three lectures, we have focused more on to the structure and as well as function of the different parts, what is present in these type of cells. So when we were discussing about the prokaryotic cell, we have discussed about the plasma membrane, we have discussed about the cell membrane, we have discussed about the plasmids and we have also discussed how you can be able to isolate the plasmids and how the plasmids can actually be able to exist in three different forms. And subsequent to that, we have also discussed about the eukaryotic cell.

When we were talking about the eukaryotic cell, we have discussed about the plant cell and as well as the animal cell, we have discussed about the differences between the plant cell and the animal cell. And we have also discussed about the different types of membrane-bound organelles what are present in the eukaryotic cell. So we discuss about the nucleus, we discuss about the mitochondria, chloroplasts, glycosomes, Golgi bodies, endoplasmic reticulum. And at the end, we have also discussed about the plasma membranes.

And while we were discussing about these cell organelles, we were also discussing about their structures, functions, and what is their contribution in regulating the different events within the cell. So in today's lecture, we are going to discuss about how you can be able to study these individual organelles. So although the membrane-bound organelles are not present within the hyperkaryotic system, we can still be able to isolate the different fractions from the prokaryotic cell and we can be able to study those fractions. Now the first question comes how we are actually going to do the fractionation of the prokaryotic cell. So within the prokaryotic cell, what are different zones where the protein is present, right? So if you see a prokaryotic cell, what you have is a cell membrane, right? You have a capsule, you have a cell membrane, then you have a cytosol and within the cytosol, you have the chromosome and other kinds of things, right? And if I show you the subcellular localizations, so you can actually have the five different fractions.

You can have the periplasmic fractions, you can have the protein which are present in the outer membrane of the cell wall, then you can have the extracellular media, you can actually have the protein which are present in the inner membrane and then you also have the protein which is present in the cytoplasm, right? This means the proteins which are present in the outer membrane, right? Protein which are present in the outer membrane, the protein which are present in the inner membrane, the protein which are present between the outer and the inner membrane, which means the periplasmic space and the protein which are present in the outer media or the protein which is present within the cytoplasm. Now, if you see the relative abundance of the proteins, now total number of protein what is present in the periplasmic fraction is 129, whereas the protein, the total number of protein which is present in the cytoplasm is 183, right? So, this means the major localization of the different proteins is either in the periplasmic space or into the cytoplasm. So, first we understand that how you can be able to get these fractions and if I want to get these fractions, what I supposed to do, right? So, this is a bacteria, right? I have to actually devise a mechanism so that I should be able to break this bacteria so that all these content should be released into the environment, right? Or into the my tube, right? And then only I can actually be able to separate them and I can actually be able to collect them separately. So, to breaking these is required are different types of disruption methods, right? So, you can actually have the different types of disruption methods to break the cells, right? So, what are different types of disruptions method and the disruption methods are specific to the prokaryotic cell or specific to the eukaryotic cell.

So, the cell disruption method.

So, cell disruption method is actually works on the basic principle of modulating the physical parameters, chemical parameters or the biological parameters. Now, if you see what we have is we have the different types of cells. We have the prokaryotic cell, eukaryotic cell, we have the plant cells, right? So, they all are very different. They all are very different in terms of the structures like for example, the bacterial cell they are very small, they are having the cell wall. So, they are actually very resistant for breakage, right? Compared to that the animal cells they do not have cell wall and their size is very large.

So, you can actually be able to use them and you can actually be able to use a delicate matter. And for the animal cell they are big, for the plant cell they are big, but they also have the cell wall. So, you can actually be able to play the three different parameters. You can actually be able to play with the physical strength such as the cell wall, right? Or you can actually be able to work with the structural organizations such as whether the plasma membrane is present or not, whether the cell wall is present or not, whether the how thick the cell wall in that or not. And then you can also use the chemical compositions.

So, you can actually be able to have the composition of the plasma membrane, you can have the composition of the cell wall and that could be different and all these can be exploited in the different types of cell disruption methods. So, considering these parameters such as physical strength, chemical composition and the structural organizations, you can also have the different types of cell disruption methods. So, what are the different cell depression methods? You can have the physical method. So, you can where you are actually going to vary the physical parameters. You can also have the chemical or the enzymatic method where you are going to use the different types of chemicals or you can also have the mechanical method which are actually going to use the different types of machines.

Now, let us first talk about the physical methods. So, in a physical method, these methods play with the physical parameter to damage the cell to release the product. So, number one is the thermolysis. This method is easy, economical and require no additional specialized equipment. It can be used only if the your protein of interest which you are interested is thermostable.

This method gives a heat shock to kill the organism and as a result, it disturb the cellular integrity without affecting the product. The effect of heat shock depends on to the ionic strength, presence of chelating agents such as DDTA and the presence of other mutilating enzymes. Then we also have the other kinds of physical methods such as osmotic shock. So, most of the mammalian cells have the plasma membrane with an active transporter to maintain the osmotic balance. Maintaining an osmotic balance is an active process with the expenditure of energy.

Prolonged exposure to the cell with a hypotonic liquid such as water causes the osmotic imbalance and ultimately causes the lysis of the cell. In this process due to the inflow of water, the cell swell and ultimately burst to release the product. According to the Hoffman's equation, osmotic pressure is directly proportional to the concentration of the solute and the temperature. So, each mammalian cell has a different susceptibility towards the osmotic shock. For example, I have given you an example of red blood cells.

So, red blood cells will life with the addition of a tiny drop of water, right. Plant or bacterial cells are more registered towards the osmotic lyses due to a thick cell wall, right. And so, what will happen is that if you have a cell, right and if you keep them under the hypotonic condition, so, for example, if you take the RBC and put it into the water, what will happen is the water molecule will start entering, right, start entering into the cell because you know that as per the principle of osmosis, the solute is actually going to enter from the lower concentration to higher concentrations. So, this is actually

going to enter and as a result, it is actually going to swell. So, it is actually going to be increase in size.

So, it will increase in size to a certain capacity. After that, what will happen is that it is actually going to break, right and when it is going to break, it is actually going to release its content and that is what exactly is happening here, right. So, osmotic shock is good for the mammalian cells, but it is not good for the plant or the cell. Then we also have another physical method which is called as sonication. So, a sonicator generates the ultraviolet waves of frequency more than the 20 kilohertz to cause the cell disruption by the cavitation.

The interaction of the ultrasound with the liquid causes the compression and decompression very rapidly. The bubbles are formed in the liquid compresses several thousand atmosphere and gives the shock wave to the cell wall or the plasma membrane to cause the cell lysis. Generation of ultrasonic waves in liquid causes a rapid change in temperature and may cause the thermal denaturation. Hence, ultrasonication media need to be cooled and very long duration should be avoided. So, this is the sonication, right and we have prepared a very small demo like where the students might have shown where students are actually going to show you how you are actually going to perform the sonication as well and what are the different types of precaution you should take.

Because as I said, you know sonication means where you are actually going to you know send the sonic waves and when the sonic waves will enter, it is actually going to cause the compression and decompression very rapidly and as a result, it is actually going to create the cavitation bubbles and these cavitation bubbles when they were actually going to interact with the cells, they are actually going to cause the lysis of the cell. But it will not happen only with the sample, it is also going to happen with your body as well and that is why you are actually going to keep this sonic probes within a cabinet and on the other hand, you are also going to take the lot of precautions so that it should not affect your sensory organs such as ears and the eyes. Now we will talk about the chemical and the enzymatic methods. So chemical enzymatic method where you are going to use the alkali treatment. So this is a harsh but effective chemical treatment to lyse the cells.

Alkali treatment cause the liquid saponification which disturb the lipid packing and affect the cell's integrity. Then you can also use the detergents and detergents are actually going to dissolve the lipid membranes and as a result, it is actually going to cause the lysis of the cell. So when you add the lipid molecules, if it is a low concentration, it is actually going to cause the pores into the cell and that process is called as the permeabilization. But if you are going to add the high quantity of the lipids,

then it is actually going to cause the complete removal of the plasma membrane. As a result, it is going to cause the solubilizations.

And then we have the enzymatic digestion. So enzymatic methods are specific, gentle and most effective but costly. Lysozyme is commercially available to treat the bacteria to release the intracellular product. In the lysosine, there are three other types of bacteriolytic enzymes, glycosidases, acetyl muralimyl-L-enulin neuradase and endopeptidase. Proprietas are also fine to be the bacteriolytic.

Each cell requires, each cell lysis require a mixture of different enzymes such as prokinesis, proteases, menaces or chitinases. Plant cell can be liked by the cellulose and the prokinesis. In most of the enzyme mediated cell lysis method, the rupture of the cell wall depends on to the osmotic pressure of the external media. In few cases, the enzymatic digestion is performed to remove the external cell wall and then in the second step, protoplasts is disrupted by the gentle solutions. So in some cases, for example, in the case of plant, you can actually be able to treat it with the enzyme.

And as a result, the plant is actually going to be without cell wall. Once the plant is without cell wall, then you can just put it into the water and it is actually going to be experienced osmotic shock and as a result, it is actually going to release the content. Then we have the mechanical methods. So we can use the different types of mechanical method. Many of these mechanical methods are very often being used in our household.

For example, you are actually going to use the mixer grinder, right. The mixer grinder is nothing but the warring blender, right. So you can actually be able to use the mixer grinder, right. So what mixer grinder is doing, it actually has a blade, right. You have seen right, they have actually blade right and these blades are running at a very high speed and as a result, what will happen is the cell is passing through these blades and when they are cells are passing through these blades, they are getting being cut, right and once they are going to be cut, they are actually going to release the content.

Then we also have the pestle mortars. Both are effective towards animal and plant tissue as well as the filamentous organisms. So you might have seen your mother or your housemaid that they are actually using the mixer grinder and as well as petal mortar for making different types of chutneys, right. That chutney is nothing but the cell extract, right. In industrial scale, cell disruption is carried out by the two different types of mechanical method, one is called bead mill and the other is called as the high pressure homogenizers.

So bead mill is bead mill disruption. The bead mill consists of a grinding cylinder with a central shaft fitted with a number of impellers which can be moved in clockwise or anticlockwise direction with the help of a motor. The grinding cylinder is filled with the beads made up of the glass, aluminia, titanium, carbide or zirconium oxide or the zirconium silicate. There is an inlet to supply the cell suspension and at the outlet to collect the sample after the process and the bead mill runs, cell experiences the shear forces between produced between the moving beads and the cells. So this is what the bead mill disruptor where you have a jacket, right and in this you are actually going to have the impellers and the impellers actually can run in clockwise direction or the antclock direction and it is actually going to have the beads.

These beads are made up of the glass or silica or titanium or the kinds of metals and when it moves these beads are also hitting the cells, right. So it is actually hitting the cells. So when the cells are getting crushed between these two beads, it is actually going to get broken down and as a result it is actually going to release the content and that release content can be collected from here, right after a certain time period. Then we have the high pressure homogenizers. So in a high pressure homogenizers, you are actually going to have the moving plates, you are actually going to have the impact rings and all that and you are actually going to feed this high pressure homogenizers with it from the feeders and when the cell is passing through these narrow gaps, because this is actually going to go up and down and as a result what will happen is that the cell will actually going to pass through with this particular narrow gap for multiple times, it is actually going to get broken down into the smaller pieces and you are actually going to change the pressure for 200 to 1000 osmosis pressure and that change in pressure is very high and that actually causes the destruction of the cells.

So this is all about the cell disruption. So once you are actually going to select the proper cell disruption method, for example, for a prokaryotic cell, you can actually be able to use the sonication, you can actually be able to use the detergents, you can actually be able to use the enzymatic method and a combination of these depending upon the product what you are trying to isolate and depending upon the stability of that particular product and so on. Now once you have done the cell disruption, you are actually going to, for example, if you are actually going to disrupt the bacterium, bacterial cell, right, you are actually going to have all the 5 fractions which we have just discussed, right. You can actually have the outer membrane, you can have the inner membrane, you can have the periplasmic fractions, you can have cytosol and you can also have the external media, right. All these 5 fractions are available and now you are actually going to devise a mechanism if you want the protein from here, whether you want the protein from here, here, here and here.

So this is what we are going to discuss. So first is how you are actually going to do the isolation of the periplasmic fractions. So what you do is you harvest the bacterial cell by centrifugation at 3000 G for 20 minutes at 4 degree Celsius, right. Remember that all these procedures, majority of the molecular biology procedures are always being performed at 4 degree Celsius except that in some specific cases where you are actually going to be asked to do it at room temperature. Then you discard the supernatant and carefully remove the last few drops of liquid with a pipette because that is actually going to contain the media. Then you gently resuspend the palette in 1 ml of the TSE buffers using a wire loops.

Then you do the 30 minute incubation on ice. So when you do the 30 minute incubation on ice, you are actually going to swell the bacterial cells. Then you transfer the cells in a micro centrifuge tube and centrifuge at 16,000 G for 30 minutes at 4 degree Celsius. Then transfer the supernatant to the new centrifuge tube and this supernatant which you are actually going to get is actually going to be the periplasmic fraction. And the periplasmic fraction you can actually be able to use.

Now the second question is how you are actually going to do the isolation of the protein from the cytosolic fraction. So what you are going to do is you are going to take the bacterial cell. What you can do is you resuspend them, resuspend in resuspension buffer. So you can actually be able to suspend them in for example phosphate buffer saline. Now you do a solication because we are not interested in getting the protein from the periplasmic fractions or other places.

So you can actually be able to, we can very easily be able to do the solication. Once you do the solications after that you are going to do a centrifugation at for example at low speed first right to 3000 G for 5 to 10 minutes at 4 degree Celsius that is actually going to give you a pellet and going to give you the supernatant. That supernatant again you are going to spin at 16000 G okay for 25 minutes at 4 degree Celsius. And again you are going to get the pellet and you are actually going to get the supernatant. The supernatant is actually the protein which are present in the cytosolic fractions.

Remember that in this pellet you are actually going to have the cell wall which means outer and inner cell wall. And whereas in this pellet is also a cell wall components right because you know that the bacteria does not contain any kind of membrane bond organelles. So basically you can actually be able to get this. From this cytosolic fractions you can be able to isolate most of the proteins. Now let us talk about the fractionation of the eukaryotic cell and remember that the fractionation of the eukaryotic cell is more complicated than the prokaryotic cell because it has a very different types of the membrane bond organelles.

So here we are discussing about the fractionation of the animal cell and as well as the plant cell. So fractionation for most of the organelles are going to be the same except that some of the fraction some of the exclusive organelles are only present in the plant cell. So for them we are going to discuss in detail differently. So for example purpose I have taken the example of how you can be able to isolate or fractionate the plasma membrane mitochondria cortisol and as well as the chloroplast from a eukaryotic cell. And if you want to do the fractionation of a eukaryotic cell you should also be able to or you should be very familiar with the different types of centrifuges which you are going to use for fractionation.

So we have the different types of centrifuge. You can have the microfuge or I will say low speed centrifuge. So it actually is going to be used for very low speed centrifugation and this is not having a low temperature and this is the high speed centrifugation. This is also a high speed centrifugation and this is a ultra centrifuge. So ultra centrifuge actually can go more than 1 lakh g RPM. So it actually can go and this is actually the ultra centrifuge rotor and this rotor is made up of a metal which is called as titanium and this is very sturdy.

And this is the same high speed centrifuge but it is actually going to be a low temperature high centrifuge tube. So these are the different types of centrifuges what we actually require if you want to do the cell fractionations. Now as far as the technique is concerned the centrifugation can be done in two different modes. It can actually be able to do a differential centrifugations or it actually can be done in a density gradient centrifugations. Differential centrifugation means that you are actually going to run the sample at a different speed and you are actually going to do the fractionation.

Similarly in a density gradient centrifugation you are actually going to change the density of the media and as a result it is actually going to separate. So we will discuss both of these aspects or different mode of the. So let us first start about the differential centrifugation. So differential centrifugation is based on the differences in the sedimentation rate of the biological molecules because of their different size, shape and density. For example you have different types of biomolecules which are actually different in terms of size, shape and density.

So what will happen is that their centrifugal power is actually going to be different and it is actually going to be affected by the solvent. It is actually going to be affected by the different types of particle size and it is actually going to be different by the other kinds of things. So what will happen is that when you do the centrifugation at a particular wave layer, particular speed that speed may not be able to provide the different enough force

to settle down the small size particle or medium size particle. So as a result the large size particle are actually going to be get pelleted down first. Medium size particle are actually going to be pelleted down second and the small size particles are actually going to be pelleted down at later stage because the small size because they are small size they require more force to get sediment.

Now let us understand this by example. For example you have the different types of molecules right. You have a iron block which is 100 kg, you require have a stone which is 30 kg, you have also an iron block which is 10 kg, you have another stone which is 10 kg, then you also have a cotton which is 8 kg and then you also have iron which is 1 kg. So what will happen is that when you do the centrifugation or when you are going to do the differential centrifugations irrespective of the weight right, irrespective of the weight the iron is actually going to get settled down first right. And then you are going to have the stone which is actually of the different density, this is actually of the highest density, this is the medium density and this is the lighter density right. So because the cotton is the lightest density it is actually going to be pelleted down at a very very high speed compared to the iron or the stone.

Same is true for the different types of biomolecule also. For example if you see the density versus sedimentation coefficient what you will going to see is that the sedimentation of the different types of organisms. For example the nucleus, mitochondria, microsomes, virus, soluble proteins all they are actually going to have the different sedimentation coefficient and as a result they are actually going to be get pelleted down at a different time points. For example the DNA is if you see the sedimentation right this is actually going to be different, RNA is going to be different, DNA is going to be different and soluble protein are actually going to be different. And this is can be exploited in a differential centrifugations to isolate in the different types of cell organelles and as well as the different types of fractions. Let's take an example of how you can be able to do the fractionation of the different organelles from organ.

So we have taken an example of the liver. So liver you know that the liver is made up of the cell which is called as hepatocytes. Apart from that it also contains the comfort cells and other kinds of blood cells but we assume that it is actually a pure hepatocyte. So the first thing what you are going to do is you are going to do the homogenization which means you are actually going to break the tissue so that it is actually going to form the single cell suspension in the first stage and then that single cells are also going to be broken down by the homogenizers. And as a result it is actually going to have the mixture of the whole cell and as well as the mixture of the different types of organelles. So imagine that these are the nucleus, endoplasmic reticulum, mitochondria and all that.

And all these organelles are different in terms of their sedimentation coefficients, in terms of their density, in terms of their size. So that actually is going to affect their sedimentations. So what will happen is that you are going to have a mixture of organelles in the beginning of the start. Now what you are going to do is first you are going to do a very low speed homogenized sedimentation. So what you are going to do is you are going to run a centrifuge at 6,600 G for 10 minutes and because this is very, very low speed you will not be able to pellet down the other material but you will be able to pellet down the nucleus, right.

Or you will be able to pellet down the unbroken cells because they are also going to be very heavy. So this is going to be a nucleus and as well as the unbroken cells. Now this portion you can actually be able to take out and put it into another tube and then you are actually going to run this at 15,000 G for 15 minutes. And when you do that it is actually going to pellet down the subsequent molecules. So subsequent to that it is actually going to pellet down the mitochondria, lysosomes and paroxosomes because they are actually going to be of heavy in nature in terms of density and in terms of size also.

But it is actually going to release the endoplasmic reticulum, it is going to leave the plasma membrane and it is also going to leave the ribosomes and other kinds of small vesicles. So this is actually going to be present here. Now if you take out these and put it into another tube and run it at a differential centrifugation for 1 lakh G for 60 minutes then what will happen is that it is actually going to pellet down the plasma membrane, it is going to pellet down the ER, it is actually going to pellet down the small vesicles. Now still it is not, it has not pellet down the ribosomes, it has not pellet down the micro macromolecules like the ribosome machinery and RNA polymerase and those kind of multimeric proteins and it also has not pellet down the virus which is very very very small right. Now what you do is you take this supernatant into another tube and run it for 3 lakhs G for 2 hours right.

So when you do that it is actually going to pellet down the ribosomes, viruses and the large macromolecules and this is what it is actually going to be left and that is called as the cytosol okay. That cytosol is only containing the monomeric proteins right. So this is the power of the differential centrifugation through which you can be able to isolate the different organelles and you can be able to test them for their characteristics and other kinds of things right. So you can actually be able to use the different types of analytical tools to characterize that you have isolated the nucleus or you have isolated the mitochondria or the phagopodia or the plastid chloroplasts and all that.

I have taken another examples of cell fractionation. The scheme is almost the same that but the source is different. So here I have taken the muscle tissue. Again you are going

to do the homogenization that is actually going to give you the mixture of the cell lysate and also going to contain the unbroken cells when you do the you know the 1000g for 10 minutes it is going to remove the contacryl ductile then you are going to take supernatant if you run it at 20000g it is actually going to separate the mitochondria when you take the supernatant and run it for 1 lakh g it is actually going to produce the crude microsomes and cytosol is different and then if you run that the stigradient centrifugation you can be able to separate out the other organelles. So this is all about the differential centrifugations. Now because these biological molecules and as well as the organelles are different in terms of density you can also be able to use the density gradient centrifugation.

So what you see here is the densities of the biological molecules are different. You have the pocratic cells, you can have the mammalian cells, organelles, proteins, DNA, RNA all these are very very different right. And what happened is that they are actually going to be get sedimentated when the sedimentation force is higher than the sedimentation rate right. So what happened is that if you actually going to have the water as a buffer right if you take a tube right and fill it with the buffer it is actually going to have the same density right. So this is actually going to have the this layer also has the same density, this layer also has the same density, this layer also has the same density.

This means it will not going to create a barrier. But what will happen is that if I take a tube and if I put the solution of different densities. So what will happen is that if the molecule is having a sedimentation coefficient which is good enough to break this barrier because you will not be able to enter inside this then you are actually going to be focused here right. Similarly if you pass this layer but you will not be able to pass this layer then you will focus here. If you pass like that that is how it is actually going to have the differential centrifugation. So what you are going to do is you prepare a density gradient centrifugation which you are going to take the tube and you can actually have the increasing density gradient or you can actually have the decreasing density gradient.

And what will happen is that you will overlay the sample right on top. So all the three samples are ready right and when you spin them you are actually going to push them in the lower side. So when they push them they will actually going to use their density as a force right. So because of that they are actually going to break the layers which are being formed and as it will go further down the density will be keep increasing keep increasing right. So it will be of lower density on the top and will be on a higher density at the bottom. And as a result it is actually going to be localized at their respective density band which means this yellow one is of a very high density and this blue ones are of low density.

And as a result they will be getting localized in that particular density zone. And now what you are going to do is you are going to take out these densities by using the different types of techniques. So you can actually be able to know do the fractionation you can actually be able to know poke the tubes from the bottom and you can be able to collect the blue cyan and as well as the yellow bottles right. Apart from that you can also be able to use different types of methods. So there are different methods of collecting the fractions from the density gradient certifications.

One is manual collection with the help of the pipette. So what you see here this is the pipette and I am sure majority of you actually know how to run how to use the pipette. And if you are not aware of then you can actually be able to follow one of my MOOCs course which is on the experimental biotechnology where I have explained and as well as given you a demo about how you can be able to use the pipettes, how you can be able to measure the liquids properly and dispense the liquids. And this is the different types of pipettes what we have. Then you can also use the automatic fraction collector for unstable gradients.

So and you can also be able to use the freeze drying technique. In a freeze drying technique what you are going to do is you are going to take this tube where you have the different zones right and then you freeze them. And after the freezing you can actually be able to collect the slices just like as we cut the slices of ice cream brick right. If you have an ice cream brick and if you have your friends in your home what you do is you cut these brick into small pieces and then it is actually going to be taken up one slice to each friend. Same is true here. You can actually be able to make the freeze then you break and you cut the slices and then each slice you collect it into a separate tube and that you are actually going to test for the presence of the different types of organelles.

So with this I would like to conclude my lecture here. In our subsequent lecture we are going to discuss some more aspects related to the biological system. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 02
Basics of Biological system (Part 2)
Lecture-06 Cellular Metabolism (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from department of bioscience and bioengineering IIT Guwahati and what we were discussing, we were discussing about the molecular biology and so far what we have discussed, we have discussed about the basics of the biological system. So, if you recall in the previous module, we have discussed about the cellular structures, we have discussed about the prokaryotic cells, eukaryotic cells and then we also discuss about the different types of organelles, what are present in the prokaryotic cell. And in today's lecture, we are going to extend our discussion about the basics of the biological system and in this context, we are going to discuss about how the cells are actually going to acquire the energy. So, in today's lecture, we are going to discuss about the cellular metabolisms and how the cells are performing the different types of metabolic reactions to acquire the energy and how it can actually be able to utilise that energy to synthesise the different types of biomolecules.

Now, the first question comes, why the cell is actually requiring the energy, whether it is a prokaryotic cell or whether it is a eukaryotic cell such as the animal cell, plant cell, fungi, it requires the energy, it requires to run the cellular metabolisms to acquire the energy, right. Because whether it is a prokaryotic cell and prokaryotic cell actually require the energy to grow, replicate and produce and you will see in the later on in this particular course that there are so many different types of cellular activities which are operating within the cell and that they are very, very crucial for maintaining the cell. For example, one of the crucial factor is the DNA replication and DNA replication is very important because it is actually required not only for synthesis of the new DNA strand, but it also requires for the repairing of the damaged DNA and repairing of the damaged DNA is important because it is actually going to protect the organisms for going for the death pathway. So, energy can be acquired from running the cellular metabolisms.

And as far as the cellular metabolism is concerned, we have two different types of cellular metabolisms. We have either the catabolic reactions or the anabolic reactions. Catabolic reactions are as I said, you know, they are the energy producing reactions. So, these are the reactions which are actually going to produce the energy. So, in this you are actually going to use the biomolecule which is responsible for producing the energy such as the carbohydrate and mostly the lipids.

These are the two major biomolecules which are being used for producing the energy. Under a very, very strong and very, very starvation conditions, the organisms can also utilize the proteins for the energy production. So, in those cases, the protein is going to be get converted into the carbohydrate and lipids and then it is actually going to run the catabolic reactions to produce energy. But that is very rare and it happens only under those conditions when you are going through starvation reactions. So, and then once you are actually going to do the catabolic reactions for the carbohydrate or lipid, you are actually going to produce the energy and that energy would be in the form of the ATP.

And this energy is actually going to be utilized for the anabolic reactions. So, this energy is actually going to utilize to drive the reaction so that you are actually going to have the synthesis of the new biomolecules. For example, if I want to synthesize a protein and you will see this when we are going to discuss about the biomolecule into the later part of this particular course that a formation of a bond is required. So, protein is made up of the amino acids. And these amino acids for example, the amino acid 1 and it is actually going to be converted going to be attached to the amino acid 2 by a bond which is called as the peptide bond.

And you know that the bond formation is always required that you are actually going to spend some amount of energy. So, when you spend the energy, you are actually going to activate the functional group what is present onto the amino acid number 1 and to amino acid 2 and that is how they are actually going to combine together and they are actually going to form a protein or peptide for example, where they are actually going to be linked by a peptide bond. So, basically what your cellular metabolism is a summation of all the reactions whether it is the catabolic reaction or the anabolic reactions. So, the anabolic reactions are required for the production of energy, whereas the anabolic reactions are required to utilize this energy for the biosynthesis because once the synthesis is done, it is actually going to contribute in terms of the growth of the organisms or the other kinds of functions. For example, it is actually going to help to produce the gametes, it is going to help the produce to give the nutrition to the daughter cells and so on.

So, these two reactions are always been under the coordination to each other and as a result, they are actually going to be responsible for the cellular health of the particular cell. So, let us start first with the catabolic reactions and we are going to start with the carbohydrate metabolisms. So, very briefly we are going to discuss about the carbohydrate metabolisms, then we will discuss about the lipid metabolisms and mostly we are going to discuss about the catabolic reactions. And then we are going to discuss about the anabolic reactions and at the end, we are also going to discuss how the you

know the cellular metabolism is taking care of the toxic products being produced during the catabolic reaction or to the anabolic reactions. So, when we talk about the carbohydrate metabolism, carbohydrate metabolism and you know that the carbohydrate metabolism is going to start once you have any food.

So, for example, if I have a food for example, if I have the rice right. So, if I have a rice in the lunch, what will happen is the rice will enter into my stomach right and then followed by the stomach, it is actually going to enter into the small intestine and from its small intestine, it is actually going to enter into the large intestine and afterwards it is actually going to be the undigested product is going to be removed from the anus right. So, food whether it is a rice right, rice is a good source of carbohydrate right, but this carbohydrate is a polymeric carbohydrate right. So, it is actually going to have the starch. Now starch you cannot put it into the catabolic reactions.

The starch has to be converted into the simple sugar such as glucose and fructose and they will enter into the catabolic reactions. So first you are going to take the rice, you are put it into the stomach, in the stomach it will actually going to start digesting right and from the small intestine, the starch is completely going to be get converted into glucose right and glucose is a monomeric sugar which is going to be ready to be get into the catabolic reactions and then this glucose is going to be absorbed by the willy and the micro-willy what is present onto the small intestine cell surface and they are actually going to be present into the blood right. So, once they are absorbed, they are going to put into the blood and from the blood, it will enter into the different organs. So, it is actually going to enter into the liver, it is actually going to enter into the muscles and so on. So brain, spleen and all that okay.

So all the organs, it is actually going to be get distributed and within the liver, muscles, spleen, brain, nervous tissues and all other, all these places, these glucose is actually going to be utilized for running the catabolic reactions and they at the end, they are going to produce the energy right. And when you have the excess amount of glucose, that glucose is going to be stored in the form of the glycogen within the different types of tissues. So that when you are going to do the starvation, that glycogen is getting converted into the glucose and that is how actually it is actually going to provide you the running force for the moment when you are not taking the nutrition from outside. So for example, when you sleep in the night, you are going to take the dinner right. But that dinner is actually going to serve the food for few hours.

After that, it is actually going to start utilizing the stored material what is present in your liver, muscles, spleen and brain. So these are the catabolic reactions what is going to be utilized right. And they will be utilized to produce the energy and that energy is

actually going to be utilized for running the normal reactions what is going to be performed by the different organs or they are actually going to be utilized for bio-acetic pathway. Now as far as a carbohydrate is concerned, it is actually going to be the central pathway for catabolism. So that is why it is actually very important that we should understand the carbohydrate metabolism.

So carbohydrate metabolism is in the central right as far as the catabolic reaction is concerned and majority of the pathways are actually getting diverged from the carbohydrate metabolism. Now, once the glucose is produced, it will enter into the different types of organs and within the different types of organs, it is actually going to first into a series of reactions which is called as the glycolysis and at the end of the glycolysis, it is actually going to produce the pyruvic acid and that pyruvic acid will enter into another cyclic reaction which is called as the Krebs cycle. So in today's lecture, we will discuss about the catabolic reactions such as the glycolysis and the Krebs cycle. So glycolysis, glycolysis is a central pathway to the carbohydrate metabolism and it is the universal pathway which is found in the prokaryote or the eukaryotic cell. It is a breakdown of a six-membered glucose into the two three-membered carbon sugar to feed the carbon to Krebs cycle in the presence of oxygen.

So you can have the two different scenarios when you have the oxygen present such as the organisms like us and it can be also be functional even if you do not have the oxygen, right. So if it is present in the presence of oxygen, the glucose is getting converted into a two three-membered carbon sugar and that is going into the Krebs cycle or it is actually going to send for the anaerobic oxidation in the absence of oxygen. So in the different types of pathogenic organisms like bacteria and other kinds of organisms, when the oxygen is limiting, it will not go to be get converted into the pyruvic acid. Instead of that, it is actually going to be get converted into the anaerobic products and that is how they will, these organisms are also going to survive. So hence, it is play a pretty crucial role for the adaptation of a living organism under the different types of stress conditions.

The glycolysis is a 10-step chemical reaction to enable the glucose for its optimal oxidations. So glycolysis is a 10-step reaction, right and these are the 10 steps, right. You have the step number 1, 2, 3, 4, 5 and in the 10th step, you are going to generate the pyruvate. So this is the step number 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10, okay. So in the step number 1, you are actually going to do the activation of the glucose molecule.

Okay, because you have to invest some energy so that the glucose will be destabilized because once you add the phosphorylated group to the glucose molecule, it is actually going to contain the very high energy and when it contains the high energy, the high energy is always making the system unstable. So phosphorylation of the glucose, so the

glucose produced after the glycogen breakdown is phosphorylated by the enzyme which is called as the glucokinase. Remember that glucokinase is only present in the liver whereas the hexokinase is present in all other tissues especially in the muscles. So in most of the organs such as brain, spleen, muscles, it is hexokinase which is the major enzyme which is going to catalyze the reaction number 1 but in the case of liver, it is actually going to be the glucokinase. So in the phosphorylation reactions, in the phosphorylation reaction, the phosphate that is the gamma phosphate group of ATP is transferred to the glucose to form the glucose 6-phosphate.

The phosphorylation reaction of glucose to produce the glucose 6-phosphate marks the molecule for the glycolysis and in this process, one molecule of ATP is utilized in the step. So once the glucose which is unphosphorylated, so this glucose is actually going to be produced from the glycogen. Remember that I talked about the stored glycogen. So once you require the energy, that glycogen is going to be broken down and it is actually going to form the glucose. This glucose can participate in the different types of reactions.

So to commit this glucose for the carbohydrate metabolism or the catabolic reactions, what you are going to do is you are going to take the carboglucose and with the help of the hexokinase or the glucokinase, it is going to be converted into glucose 6-phosphate. Once you generate the glucose 6-phosphate, there is a big difference. This is the neutral molecule. There is no charge on the glucose molecule. Whereas once you generate the glucose 6-phosphate, this is actually going to be the negatively charged molecule.

And once you generate a negatively charged molecule, you are actually going to trap the molecule within the cell. Because a charged molecule cannot be freely available to go out of the cell, because the charged molecule, the movement of a charged molecule from the cell requires the energy. This means that glucose 6-phosphate is going to be entrapped within the cell and then it is actually going to be committed for no other reaction, but that glycolysis and then it will actually go to do the reaction number 2. Now, in the step number 2, there will be a conversion of glucose 6-phosphate to the fructose 6-phosphate because you are going to have the isomerization reaction. So, in the step number 2, the glucose 6-phosphate that you have generated from the glucose is going to be converted into fructose 6-phosphate and the enzyme that is going to catalyze this is called as a phosphofructo isomerase.

Now, in the step number 3, another series of oxidation or another series of phosphorylation is going to take place in the step number 3. So, in the step number 3, you are going to have the phosphofructo kinase catalyzing the another round of phosphorylation. So, you have first phosphorylation here and you have the second

phosphorylation here and as a result, you are going to generate a molecule which is called as fructose 1, 6 bisphosphate and this molecule is a very high energy molecule and very unstable molecule. So, in the step number 3, the sugar is further phosphorylated at the carbon number 1 to produce the fructose 1, 6 bisphosphate by the action of an enzyme which is called as phosphofructo kinase. In the phosphorylation reactions, the phosphate that is a gamma phosphate group of ATP is transferred to the phosphorylated sugar to form the fructose 1, 6 bisphosphate.

One molecule of ATP is utilized in this step. So, remember that we have utilized one ATP here and one ATP in the reaction number 3. Now once the fructose 1, 6 bisphosphate is generated, which is actually we have very, very, very unstable molecule, it is actually going to be act by the aldolase in the reaction number 4. So, in the reaction number 4, the aldolase is actually going to break or it is actually going to break the molecule into the two different molecules, the glyceraldehyde 3 phosphate or the dihydroxyacetone phosphate. So, in the step number 4, there will be a cleavage of fructose 1, 6 bisphosphate and this step is catalyzed by an enzyme which is called as the aldolase or fructose 1, 6 bisphosphate and aldolase to generate the glyceraldehyde 3 phosphate which is called the aldose and the dihydroxyacetone phosphate which is called as the ketose.

So, there will be a cleavage of this high energy bond. Remember that until this, you have the 6 membered carbon, 6 membered ring. Now at this stage, it is actually going to be get converted into the 3 membered rings and that is a cleavage of the sugar molecule. Now in the step 5, this is the step number 4, you are going to have the isomerization reactions and the conversion of the dihydroxyacetone phosphate to the glyceraldehyde 3 phosphate. So, interconversion of the triose phosphates, so 3 carbon sugar formed in the step 4 undergoes the internal convergence as the glyceraldehyde 3 phosphate can readily be entered into the next step.

The ketose generated in the step 4 is reversibly converted into the glyceraldehyde 3 phosphate by the triose 3 phosphate isomerase. So this is also a very, very important enzyme because it is actually going to convert the dihydroxyacetone phosphate into the glyceraldehyde 3 phosphate. And now afterwards, so it is actually going to generate the 2 molecules of glyceraldehyde 3 phosphate. Remember that from fructose 1, 6 base phosphate, aldolase is actually going to break this into the glyceraldehyde 3 phosphate and the dihydroxyacetone phosphate. So it is actually going to generate the 1 molecule of this and 1 molecule of this, but with the action of phosphotriose isomerase, this molecule is getting converted into this and as a result of this, the 2 molecule of glyceraldehyde 3 phosphate is going to be generated from the fructose 1, 6 base phosphate.

Now in the step number 6, this is the first time when you are actually going to see a generation of the ATP. So you are going to see the generation of the reducing equivalence and this reducing equivalence are actually going to produce the ATP. So in the step 6, what you are going to do, so this is the step number 6, what we are going to do is you are going to see the dehydrogenase reactions and there will be a generation of the reducing equivalence. So in the step 6, the glyceraldehyde 3 phosphate, the glyceraldehyde 3 phosphate is going to be get converted into the 1, 3 base phospho-desirate and in this step, the 1 molecule of NADH is produced after the oxidation of the aldehyde group of the glyceraldehyde 3 phosphate with the help of the enzyme glyceraldehyde 3 phosphate dehydrogenase. This enzyme is very important for the many types of the therapeutic applications such as generation of the different types of drugs and other kinds of thing because this is the enzyme which is actually going to be first time going to produce the reducing equivalent and these reducing equivalence when they will put into the electron transport chain, they are actually going to produce the ATP.

So that is why if you mutate or if you inhibit this particular enzyme, you are actually going to destroy the glycolysis and you are also going to block the production of the energy. Even in those organism where the oxygen in the absence of oxygen, the glycolysis is going to be keep running and keep giving them the enough energy so that they can be able to survive under the stress conditions. Now once you generated the 1, 3 base phospho-glycerate from the glyceraldehyde 3 phosphate, it will be converted into the next reaction and the next reaction is the seventh reaction. So in the step 7, in this step, the phosphate group from the 1, 3 base phospho-glycerate is removed by the phosphoglycerate kinase with an acyl group transferred to the ADP to generate the ATP molecule. So in this is the first step where you are directly going to see a generation of the ATP molecule and this is the enzyme which is actually responsible for generation of the first ATP molecule.

And 1, 3 base phosphoglycerate is going to be get converted into 3 phosphoglycerate and the phosphate what is present on the carbon 1 is actually going to be taken up by the ADP molecule and as a result, it is actually going to generate the ADP molecule. So this is the step number 7 which is actually the step which is going to generate the energy first time. Remember that in the step number 6 also you have generated the energy but that is indirect energy. It is actually going to get into the electron transport chain and then it is actually going to produce the ATP. But here directly you are going to get the ATP molecules.

Now from the 3 phosphoglycerate, there will be isomerization reactions and it is actually going to get converted into the 2 phosphoglycerate. So in the step number 8, you are going to have the conversion of the 3 phosphoglycerate to the 2

phosphoglycerate which means there will be a change of the position of the phosphate group within the molecule and this reaction is going to be catalyzed by an enzyme which is called as a phosphoglycerate mutase and it is actually going to form the 2 phosphoglycerate. Now in the step number 9, there will be a dehydration of the 2 phosphoglycerate to phosphoenol pyruvate. The enzyme enolase catalyzes the dehydration reaction to produce the phosphoenol pyruvate, a compound with a high phosphoryl group transfer potential. So now from the 2 phosphoglycerate, the enolase is actually going to remove the one molecule of the water and as a result, it is actually going to form the phosphoenol pyruvate.

And from the step number 10, which is the last step of the reaction, so this is the step number 9, which is the step number 10, the phosphoenol pyruvate is actually going to give up another phosphate and as a result, it is actually going to generate the pyruvate and the enzyme phosphoryl pyruvate kinase and here again, you are actually going to produce the direct energy which means it is actually going to produce the instant energy and it is actually going to produce the ATP. So the first time, you have produced the energy here and the second time, you are going to produce the energy here. And in the step number 10, the phosphate group from the phosphoenol pyruvate is transferred by the pyruvate kinase with an acyl group, phosphate group transferred to the ATP to generate the ATP molecule. Now this is the glycolysis ATP balance sheet and what you see here is that I have given you that how much amount of investment and what will be the production. So in the investment, remember that in the step number 1 to 4, there is an investment of two ATP molecule because you are utilizing the ATP molecule in the step number 1 and in the step number 3.

Now, once you have invested the two ATP molecule because you have activated with the step number 1, you have activated the molecule and you have phosphorylated the glucose so that it will be committed for the glycolysis and in the step number 2, you have phosphorylated fructose 6 phosphate so that it will be going to produce the fructose 1, 6 bisphosphate and it is actually going to be ready for the cleavage reactions. And fructose 1, 6 bisphosphate is a very, very high energy unstable molecule. So once you generated the unstable molecule, it will actually going to go into the downstream reaction. So in the step number 6, the ATP is actually going to be produced, NADH is actually going to be produced and that NADH when it goes into the electron transport chain, it is actually going to give you the ATP molecule. Then in the step number 7, there will be a generation of ATP when you are leaving the one ATP molecule to the one phosphate groups to the ADP and in the step number 10, which is the final step phosphoenol pyruvate is also giving one phosphate molecules to the ADP and that is how you are actually going to have the two different types of ADP molecule.

And remember that after the step number 5, you have the cleavage of fructose 1, 6 bisphosphate to the glyceraldehydes 3 phosphate and dihydrogenes phosphate. So you are actually going to have two molecules of the glyceraldehydes 3 phosphate. So the one molecule when going to process, it is going to produce the one molecule of NADH, one molecule of ATP and one molecule of ATP in this. But since you have two molecules of NADH, two molecules of the d33 phosphate, it is actually going to produce the two molecules of NADH and when the two molecules of NADH is going to process, it is going to generate the six molecules of ATP. And in the step number 7 and 10, it is also going to generate the two molecules.

So total T, what you are going to see here is that this is the final balance sheet. So 6 is from the NADH, 2 is from the reaction number 6 and 2 is from the reaction number 10 and these are the two reactions, two ATP what you have actually invested. So at the end, you are actually going to have the 8 ATP molecules. So at the end of the glycolysis, if you have the oxygen present, one molecule of glucose is actually going to give you the 8 ATP molecules. Now, if it is a reaction, it is actually going to be regulated by many methods.

So one of the major method is that you are actually going to regulate the level of glucose and that is always being done by the different types of hormones. You know that the different types of hormones are regulating the concentration of the glucose within the blood and outside within the cell also. And one of the such hormones is called as insulin hormone. And insulin actually binds to a septum which is called as the insulin septum and these are the protein tyrosine based receptors and they will actually going to you know drive the reaction inside the cell in such a way that it is actually going to down regulate the glucose. So it will actually going to enhance the uptake of the glucose.

So what happened is that when you have the insulin binding to the insulin receptor, it is actually triggering the opening of the glucose transporters. So you have a glucose transporter and they are actually going to increase the entry of the glucose inside the cell and once they are actually going to enter inside the cell, so they will be going to take up the glucose, they will be going to take inside and then they will recycle and go back. So in the step number one, the insulin will bind to the receptor into the cell membrane and activated receptor promote the recruitment of glucose transporter from the intracellular pool to the cell membrane. So once that happens, you are actually going to have a very high concentration of the glucose transporters such as GLUT3 and GLUT4 and they will actually going to enhance the uptake of glucose from the bloodstream. And once it enters and suppose the glucose is less, then what will happen is that these transporters are actually going to be taken up into the intracellular vesicles.

And by doing this, it is actually going to regulate the level of glucose into the blood. Apart from this, the glycolysis can also be regulated at the level of the feedback mechanism and as well as the covalent modifications. So this is the example of the covalent modification and this is the example of the allosteric regulations. And I am not going to discuss in detail about this because this course is more about the molecular biology, but what will happen is that in the case of the covalent modifications, the pyruvate kinase, which is actually the enzyme that is catalyzing the 10th reactions can be present in two different forms. It can be a phosphorylated form or it can be a dephosphorylated form.

The dephosphorylated or I will say the native enzyme is actually very active, but when it gets phosphorylated, it becomes less active. So because of this, it can actually be able to get modulated by the different types of parameters. For example, if there will be a low blood glucose, it is actually going to drive the reaction in such a way that it is actually going to convert the dephosphorylated pyruvate kinase to the phosphorylated pyruvate kinase and so on. Apart from that, you are also going to have the modulations either by the fructose monistic bisphosphate, the level of phosphor and the ATP and alanine. So ATP if there is a sufficient quantity of ATP what is present inside the cell, it is actually going to down regulate the activity of these enzymes.

Whereas if the level of fructose monistic bisphosphate is very high, it is actually going to increase the activity of this particular enzyme. Same is true when you are talking about the allosteric regulation. So here also you have the many types of allosteric regulators. So phosphor-fructokinase is actually an enzyme which is going to be allosterically be regulated by the fructose 2, 6 bisphosphate. So what is mean by the allosteric regulation is that the molecule will not going to bind to the active site, but it will bind to a allosteric site and because of that either it will increase the activity of that particular enzyme or it will actually going to decrease the activity of that enzyme.

And either of these ways you are actually going to see or you are going to be able to regulate the enzyme activity and at the end you are going to regulate the glycolysis. Now from the glycolysis the glucose is going to be get converted into the pyruvate. Now this pyruvate will enter into another chain, another reaction which is called as the Krebs cycle and Krebs cycle is a chain reaction cycle. So the Krebs cycle as the name suggests the Krebs is a name of the scientist and the Krebs cycle is discovered by Professor Hans Kep and it has all sugar intermediate with the 3 carbons. Remember that in the glycolysis we have started with the 6 carbon and then it will enter into the 3 carbon whereas in this case all the carbohydrates are of 3 carbon sugar.

It is also known as the tricharmoxic acid or the citric acid cycle. In higher eukaryotes

the Krebs cycle operates inside the mitopondyl stroma with the different enzyme. In the presence of oxygen the pyruvate formed during glycolysis enter into the Krebs cycle for further oxidation to produce the energy. So what we have is we have the pyruvate right. So this pyruvate is coming from the glucose from the glycolysis right.

Now this pyruvate is going to be entered into the Krebs cycle. So it will get converted into the acetyl CoA and pyruvate is actually going to converted into acetyl CoA with the enzyme which is called as pyruvate dehydrogenase complex. This is a multi-manic enzyme complex and it requires the different types of cofactors like TPP, lipoate and in this process one molecule of NADH is actually going to be produced. Now acetyl CoA is actually going to enter into the Krebs cycle. So this is actually the Krebs cycle right and the acetyl CoA is going to be combined with the water and it is actually going to form the citrate and the enzyme is citrate synthase.

So this is the reaction number 1 okay. Now from the citrate you are going to have the two reactions of the dehydration reactions. So in the first step of dehydration when the first molecule of water is going to be removed by the enzyme aconitase it is actually going to form the cis aconitase and from the cis aconitase when there will be another round of removal of water it is actually going to form the isocitrate okay. So this is the reaction number 2, this is the reaction number 3 and now in the reaction number 4 the isocitrate is going to be get converted into oxaloacetate and the one molecule of NADH is actually going to be reduced and the enzyme which is going to catalyze this reaction is called as the isocitrate dehydrogenase. Now from the oxaloacetate it is going to be there will be a decarboxylation reactions and as a result there will be a removal of carbon dioxide and that is how it is actually going to form the alpha-ketoglutarate and the enzyme is isofumarate dehydrogenase. And from the alpha-ketoglutarate there will be a generation of the NADH molecule and there will be removal of the decarboxylation reaction.

So it is going to produce the one molecule of carbon dioxide and there will be a generation of NADH molecule and the enzyme which is going to catalyze the conversion of the alpha-ketoglutarate to succinyl CoA is the alpha-ketoglutarate dehydrogenase. Now from the succinyl CoA you are actually going to generate the succinate and in this process it is actually going to produce the one molecule of GTP. Remember that the GTP is of the same energy as the ATP okay. And then this enzyme this reaction is going to be catalyzed by an enzyme which is called as succinate thiokinase. Now from succinate it is actually going to form the fumarate and the enzyme is succinate dehydrogenase and in this process the one molecule of FADH₂ is going to be produced.

And from the fumarate it is actually going to form the malate and the enzyme is called

as a fumarate and then will be a removal of water right from the malate to generate the fumarate. So there will be hydrogen reactions and from the malate it is actually going to form the oxaloacetate and in this process also there will be a generation of the NADH molecule and the enzyme which is going to catalyze this reaction is called as malate dehydrogenase and again the oxaloacetate is going to combine with the acetyl CoA from the pyruvate to form the citrate and that is how it is actually going to complete the reactions. So this is the step number 4, this is step number 5, this is the step number 6, this is step number 7, 8, 9 and this is the step number 10 okay. So by doing this cyclic reactions you are actually going to be utilized the one glucose molecule completely and you are going to oxidize that into the form of the ATP and NADH and as a result you are going to produce a very high quantity of energy especially when you are actually going to have the enormous amount of oxygen present so that you can be able to run the electron transfer chain optimally.

So let us see how much energy you are going to produce. So there is no investment as far as the phase cycle is concerned right because you are not going to invest any ATP molecule. You have already invested ATP molecule if you are talking about the glycolysis but once you activated the glucose molecule for the carbohydrate catabolic reactions then it is fine. So in the step number 1 there will be a production of a style CoA right when the pyruvate is getting converted into a style CoA and that is how it is actually going. So there is a one generation of NADH molecule and the NADH molecule is going to give you the 3 ATP molecule. Then in the step number 3 there will be a generation of the alpha-thryl-thryl-butyrate and then also you are going to have G1 molecule of NADH.

So here you have one molecule of NADH, here also you have one molecule of NADH. Then in the step number 4 there will be a generation of succinyl CoA, there also you are going to have the NADH molecule right. Then you also have the generation of GTP. So GTP is also having the same energy as the ATP so there will be one ATP molecule which is going to be produced and then in the step number 6 there will be a generation of slumarate. So there will be a generation of FADH₂ rather than NADH and it is actually going to give you the 2 molecules. So here you are going to have the FADH and the step number then 8 there will be a generation of oxaloacetate.

So that also is going to give you the one molecule of NADH okay. So at the end what you see here is this is the net balance of the oxidation of one pyruvate molecule and it is actually going to give you the 15 ATP molecule. And since from one glucose molecule you are producing the 2 pyruvate molecule so it is actually going to generate the 30 ATP molecule right. Now because so at the end if you talk about the glucose and if there is a ample amount of oxygen present what will happen is that with the help of the glycolysis

it is actually going to produce 8 ATP molecule and with the help of the grape cycle it is actually going to produce the 30 ATP molecule. And at the end from the one glucose molecule you are going to produce the 38 ATP molecule that only when you are actually having the oxygen present. If there is no oxygen present then the production of ATP from the glycolysis and as well as the production of ATP from the grape cycle is actually going to be reduced.

Because majority of these NADH molecule or the FADH molecule will not go into the grape cycle for the oxidation and as a result they will not go to produce any energy if the oxygen is not produced and same is true for the glycolysis. So, here there is a question comes what would be the oxygen what would be the amount of ATP produced when you do not have the oxygen that you are actually going to figure out and you can have to tell me. Now how you are going to regulate the grape cycle so regulation of the grape cycle can be done at the four level. One is you can actually have the conversion of pyruvate into acetyl-CoA is the first step which allows the entry of sugar moiety into the grape cycle and the pyruvate dehydrogenate complex is allosterically inhibited by the high ratio of ATP to ADP, NADH and acetyl-CoA which means if you have high quantity of energy whether in the form of the ATP or whether in the form of the reducing equivalent then you are actually going to allosterically go to reduce the activity of the pyruvate dehydrogenate complex. So if you have the ATP you have NADH you have acetyl-CoA or if you have enough quantity of fatty acids then you are not going to run the carbohydrate metabolism then you are actually going to take the fatty acid and directly enter into the grape cycle and run it.

On the other hand if you have the very high quantity of NADH, FAD plus, acetyl-CoA or calcium then you are actually going to increase the activity of this activity and the more of the pyruvate is getting converted into the styrofoam because it is actually allosterically going to enhance the activity of the pyruvate dehydrogenate complex. Then the first reaction of the phase cycle is catalyzed by the citrate synthase is inhibited by the high level of NADH, ATP and acetyl-CoA. So the first reaction is also going to be modulated by the presence of the ATP, NADH, styrofoam or fatty acids. The same logic if you have a high quantity of energy then you would not like to run the cycle. Then we have the isolated dehydrogenase which is also going to be regulated by the ATP and NADH whereas in the case of ADP and calcium which is actually going to increase the activity.

And then we have the alpha-2-proteinate which is actually going to be inhibited by the succinyl-CoA and the high level NADH whereas the calcium is stimulating the system. So this is what is given here. Now cycle cycle is a central metabolic pathway. As I said carbohydrate metabolism is a central metabolic pathway. And that is why it

communicates with the many types of metabolic pathways so that they can be able to make the good coordination.

For example you should not, you do not want that there will be enhanced production of the isocitrate. And on the other hand if one reaction requires the isocitrate for its own, you know for as a reactant right then you should take the isocitrate from here and put it into that. So that is why the TCA cycle is a central metabolic pathway and it actually requires the metabolites for the different types of the other metabolic pathway as well. And that is why what you see here is that the TCA cycle or the tricarboxylic acid cycle is having the different types of intermediate. For example it has citrate, it has the alpha-2-protonate, succinyl-CoA, malate and the oxaloacetate.

And what you see here is that if the citrate, citrate is actually communicating with the fatty acid and steroid because citrate can be used in that particular biosynthetic pathway. Same is true for the alpha-glutobutyrate, it can actually be able to communicate with the synthesis of the amino acids like glutamate and once the glutamate is formed it can actually be able to generate the glutamine, proline, arginine. All these you are actually going to see when we are going to discuss about the anaerobic oscillations, when we are going to discuss about the anabolic reactions. Same is true from here also how the oxaloacetate is communicating with the phosphoenol pyruvate, glycine, serine and all that. And carbohydrate and fat metabolism is also very actively interacting with the different types of intermediate present in the TCS cycle.

And the Krebs cycle because it is a central metabolic pathway, it is a master regulator of metabolism because it can regulate not only the carbohydrate metabolism, but also the metabolism of the other metabolic pathways such as fatty acid biosynthesis pathway, fatty acid oxidations, protein synthesis and the nucleotide synthesis. So, this is all about the catabolic reactions of the carbohydrate metabolism. In our subsequent lectures, we are going to discuss about the catabolic reactions of the fatty acids and then we are going to move on to discuss about the anabolic reactions and how the energy what you are going to generate into the catabolic reaction is going to be utilized into the anabolic reaction for the synthesis of the biomolecules. So, with this, I would like to conclude my lecture here. In our subsequent lecture, we are going to discuss more about the catabolic reactions of the lipids and as well as the anabolic reactions of the other biomolecules. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 02
Basics of Biological system (Part 2)
Lecture-07 Cellular Metabolism (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati and what we were discussing, we were discussing about the cellular metabolism and in this context in the previous lecture, we have discussed about the carbohydrate metabolism. So, we have discussed about the how the glucose is being phosphorylated within the glycolysis and that is how that particular glucose is going to be committed for the glycolysis and further you know further degradation into the crave cycle, so that it will be able to produce the energy and this energy what it is going to be produced is in the form of ATP or in the form of reducing equivalence and this energy is going to be utilized for many of the anabolic reactions. So, apart from the carbohydrate, we also have another molecule which is called lipids and these lipids are also been the major source of energy within the cell. So, in today's lecture, we are first going to discuss about the catabolic reactions of the lipid metabolic metabolisms and how it is actually generating the energy and subsequent to that we are also going to discuss about the anabolic reactions and at the end of this lecture, we are also going to discuss about the how the body is managing the different types of metabolic byproducts. So, as we have discussed in the past also in the previous lecture also that when you talk about the cellular metabolisms, you can have the two different types of metabolism, one is catabolism where you are actually going to have the energy producing reactions and within the catabolism, you are actually going to have the two different types of biomolecules, you can have the carbohydrate and or you can have the lipids.

Lipids are actually going to be processed within the glycolysis followed by the grape cycle whereas, the lipids are actually going to be processed under the beta oxidation and ultimately, both of these processes are going to produce large quantity of energy in the form of ATP slash NADH okay and this energy is actually going to be utilized into the anabolic reaction where you are going to have the biostatic pathway. So, within the anabolic reactions, we are going to discuss about the protein synthesis or amino acid synthesis and at the end, we are also going to discuss when you are going to do the catabolic reactions or the anabolic reactions, they both are actually going to produce the waste material right and this waste material also need to go through with some of the metabolic pathways so that it can be detoxifies. So, let us start with the beta oxidations. So beta oxidation as the name suggests is actually a metabolic pathway which is required for the lipids and you know the lipid is a chain, is a carbon, is made up of the fatty acids

plus glycerol right and this fatty acid part is actually going to be utilized into the beta oxidation to produce the energy and how we are going to get the fatty acids right.

So, when we consume the lipids into the food for example, if we take the food right for example, if we take the pizza right. So, pizza is actually going to have the lipids right and that lipid is actually going to be digested by the enzyme which is called lipase and that is how it is actually going to generate the fatty acid and then the fatty acid is going to be absorbed by the small intestine. And ultimately it is going to be transported to the liver and it is also going to be transported to the other body parts and it is actually going to be used for the beta oxidation to generate the energy. Now before we get into the beta oxidations, we have to first very briefly see how all these things are actually working. So what you are going to do is when you are going to take the fat into the food, fat will get into the stomach, so there is no digestion of the fat into the stomach right and then it will enter into the small intestine.

When it enters into the small intestine, the bile salt which is going to be released from the gallbladder is actually going to emulsify the fat. So what is meant by the emulsification is that it is actually going to make the fat molecules more polar in nature so that it will get dissolved into the aqueous environment and so that it will actually go to have the proper action of the different enzymes. So in this process of the emulsification, it is actually going to form the micelles and the fatty acids and then it is actually going to present into the small intestine where it is actually going to be digested and that is how it is actually going to produce the fatty acids. These fatty acids are then going to be taken up into the blood and that is how it is actually going to form the chylomicrons. And then these chylomicrons are actually going to transfer to the capillaries of the blood to the different tissues.

So it is actually going to travel in the form of a chylomicrons and these chylomicrons are going to be targeted to the different organs depending upon the type of protein what is present onto the cell surface, onto the surface of these chylomicrons and that is how they are actually going to deliver the fat to the brain, muscles, livers and also on. And the fatty acids which are entered into the myocytes and adipocytes where they undergo the degradation or the beta-oxidations. And once it is going to go through the beta-oxidation, it is actually going to produce the carbon dioxide, it is actually going to produce the ATP and it is also going to produce the reducing equivalents. Now, why there is a need to have the beta-oxidations? So beta-oxidation is the sequential removal of the two carbon fragments from the carboxyl end of the fatty acids. During the process, the acetyl CoA is going to form as the bond between alpha and beta carbon atoms are broken.

It is named so because the beta carbon of the fatty acid is oxidized and the process occurs inside the mitochondria. So beta-oxidation occurs within the mitochondria and why it is called beta-oxidation? Because the bond between the alpha and beta chain is actually going to be broken. So this is actually the acid part and this is the alpha carbon, this is the beta carbon. So what has happened is that in a cyclic reactions, the bond between the alpha and beta is actually going to be broken down. And that is how this portion is actually going to be released and it is actually going to form the acetyl CoA.

And this portion is then going to be transported into the crepe cycle and it is going to be oxidized into the crepe cycle. Remember that the acetyl CoA is actually going to be combined within the crepe cycle and it is actually going to form the citric acid. And that is how it is actually going to enter into the crepe cycle. So beta-oxidation is actually the reaction which are going to lead to the breakdown of this particular fragment and in this process also it is actually going to generate some energy. So you can imagine that if you have a carbon of pentadecanoic acid, it is actually going to have this kind of breakage after every two carbons.

So it is actually going to have a breakage here, it is going to have a breakage here, it is going to have a breakage here like that. And all these two chain carbons are actually going to be get converted into the acetyl CoA and then these acetyl CoA will be transported or will be a part of the crepe cycle and that is how it is actually going to produce the energy. So beta-oxidation is that it is actually going to produce the acetyl CoA from the long chain fatty acids and then acetyl CoA will enter into the crepe cycle. So this process also requires the multiple steps. First of the step is that it actually requires the activation of the lipid molecules and then the transported of the lipid molecules within the mitochondria.

Remember that the beta-oxidation will occur inside the mitochondria compared to that the carbohydrate metabolism starts from the cytosol. So the fatty acids are activation and the transportation to the mitochondria. Enzymes for the beta-oxidations are located in the mitochondrial matrix which means the liquid part of the mitochondria. The fatty acids with chain length greater than 14 cannot cross the mitochondrial membrane as such. Therefore, they first undergo activation and then transportation aided by the three enzymatic reactions.

Once the fatty acid reach the target cell, their activation takes place in the cytosol. So fatty acid activation is a ATP dependent acetylation reaction in which the fatty acid is activated by the coenzyme A and ATP to form the fatty acyl CoA with the help of an enzyme which is called acyl CoA synthetase or it is also called acyl CoA ligase or acyl CoA thiophenase. Thus acyl CoA synthetase catalyzes the formation of a thioester

linkage between the carboxyl group of the fatty acid and the thiol group of the coenzyme A to yield the molecule which is called as fatty acid fatty acyl CoA. So this is the fatty acid and this group is going to be combined with this group and that is how it is actually going to form the fatty acyl CoA and in this reaction the ATP is actually going to be consumed because you know that it is forming a high energy bond. So this is actually going to be linked to the CoA and that is how it is going to form the fatty acyl CoA and the energy what is present in the ATP is actually going to be consumed.

Now once you have generated the fatty acyl CoA you are actually going to commit this particular lipid for beta oxidation which means you have already invested some energy or you have already invested some amount of energy into this so that you can be able to send this lipid for the beta oxidation inside the mitochondria and then it is actually going to be you know generate more amount of energy. Now the next step is the transportation into the mitochondria. So mitochondrial inner membrane is impermeable to almost all the fatty acid CoA and molecule which are transported to the mitochondrial matrix by the carnitine shuttle. So you can have a carnitine acyltransferase shuttle where you have a protein which is called as carnitine and this protein is actually going to bind the acyl coenzyme this and that is how it is actually going to help into the transports. So you have the inner membrane, you have the outer membrane, this is the mitochondria and then how it is actually going to help is that each fatty acid coenzyme A is converted into the fatty acyl carnitine derivative in a reaction named transesterification by the enzyme carnitine acyltransferase A1 which is present in the outer membrane of the mitochondria.

So in the outer membrane of the mitochondria you have an enzyme which is called as carnitine acyltransferase 1 and it is actually going to join the carnitine to the incoming fatty acids. So the derivative is this derivative is located to the mitochondrial matrix by the acyl carnitine carnitine translocase which is present in the inner membrane of the mitochondria. This means once this is going to be formed it is actually going to be taken up into the inner membrane and then you are actually going to have the acyl carnitine translocase. This is actually going to allow the entry of this carnitine conjugated fatty acid molecules and once it enter into this then you are going to have the carnitine acyltransferase 2 which is actually going to remove. So once the fatty acid is regenerated via the carnitine acyltransferase 2 located onto the matrix side of the inner mitochondria, carnitine is transported.

So this carnitine protein is again transported back to the outer membrane and then it will be available for making a complex with the other molecule of the. So carnitine is transported back into the inner mitochondrial space via the acyl carnitine transporters which is then ready for the participation into the other reaction of the activating the fatty acids. So these are the this is the carnitine acyltransferase shuttle and it is actually going

to help in the transportation of these acetylated lipid molecules. Then we have the stages of the beta oxidation. So you can have the three stages.

Stage 1 you can have the beta oxidation, stage 2 oxidation of acetyl-CoA and stage 3 that is called as the oxidative phosphorylation. So in the stage 1 the long chain fatty acid is oxidized to yield the acetyl residue in the form of acetyl-CoA which is known as the beta oxidation. Then in the stage 2 the oxidation of the acetyl-CoA produced from the oxidation of the fatty acid is further oxidized to carbon dioxide via a cycle to yield the reducing balances. And then the stage 3 is that whatever the reducing equivalents are being generated they will be going to be electron derived from the oxidation of the stage 1 and 2 passes to the oxygen via the mitochondrial spectra chain for ATP synthesis by the oxidative phosphorylation. This means this is the stage 1 where the long chain fatty acid is actually going to be broken down after every 2 carbons.

And then it is actually going to generate the acetyl-CoA. So for example in this case it is actually a 16 membered carbon fatty acid. So it is actually going to generate the 8 acetyl-CoA. And then all these 8 acetyl-CoA is going to enter into the 30-grade cycle and it is actually going to generate the 16 carbon dioxide molecule and at the end it is actually going to generate the NADH and FADH₂ along with that it is going to generate the ATP, GTP and all that. And then in the stage 3 NADH and FADH₂ is actually going to enter into the oxidative phosphorylation and that is how it is actually going to generate the large quantity of ATP.

So let us talk about the stage 1 which is the beta-oxidations. So this is a stage 1, all the reactions of the stage 1. So once the fatty acid-CoA molecules are exported to the mitochondrial matrix, they are subjected to the repeated 4-step process. Each time the chain length reduces by the 2 carbon till the final product is the cycle itself. For example, if you start with the permeth oil, it is first going to break this bond and then it is actually going to break subsequent to that.

So there are multiple steps in the beta-oxidations. So first step is the oxidations. The first reaction is catalyzed by the 3 isozymes of acyl-CoA dehydrogenase. So it is a flavoprotein with FADH as the prosthetic group. The electrons extracted from the fatty acid-CoA are transferred to the FAD and a reduced form of dehydrogenase immediately imparts its electron to an electron carrier of the mitochondrial respiratory chain which is an electron transferring flavoprotein.

The reaction is analogous to the succeeded dehydrogenase reaction in the cystic acid cycle where FADH act as an electron receptor. So in the step 1, the acyl-CoA dehydrogenase is actually going to participate and it is going to oxidize the carbon.

Then in the step 2, there will be a hydrolysis. So in the second step of beta-oxidation cycle, the water is added to the double bond of the trans-enoyles-CoA to form the beta-hydroxyl acyl-CoA. The reaction is catalyzed by an enzyme which is called as the enoyl-CoA hydratase and which is similar to the reaction performed by the fumarate enzyme into the acetylated cycle.

And then the third step is the oxidation. In the third step, the beta-hydroxyl acyl-CoA undergoes dehydrogenation. So synthesize the beta-ketoacyl-CoA via the enzyme known as the beta-hydroxylacyl-CoA dehydrogenase. Here the NAD plus act as an electron acceptor. The NADH formed in the above reaction transferred its electron to the NADH dehydrogenase and electron carrier of the spectra chain.

So what will happen is that by the end of these beta-oxidations, it is actually going to produce the acetyl-CoA. Then we have the stage step 4 where you are going to have the thiolizes. So in the final reaction of the beta-oxidations cycle, the beta-ketoacyl-CoA is cleaved by the reaction with the thiol group of the coenzyme A to yield an acetyl-CoA molecule and a coenzyme A thioester of the fatty acid. Shortened by the two carbon atoms, the reaction is performed by the enzyme acyl-CoA-acyltransferase. So, for example, if we start with a fatty acid which is C16, the product is going to be undergo for the 8 oxidation, the product after 1 beta-oxidation will be C14.

So after 1 beta-oxidation, it is going to be C14. This means the 2 carbon which comes out is actually going to produce the acetyl-CoA. This is the acetyl-CoA which is going to be produced and the C14 carbon what is left is actually going to be go through these reactions again. This means it will go again, then there will be acetylation and all that. So this means it will continue till you are actually going to have the C2 on this side and you are going to have the acetyl-CoA.

So it is actually going to give you the acetyl-CoA at the end. This means if you start with the C16, what you are going to get is you are going to get 8 acetyl-CoA enzyme, acyl-CoA molecule which will enter into the phase cycle and on the top it is also going to produce the NADH and it is also going to produce the FADH and it is going to produce the ATP if required. So and all these 8 molecules will enter into the phase cycle. Now in the stage 2 of oxidation of acetyl-CoA, considering the permutoyl-CoA C16, 1 beta oxidation will give you the myristyl-CoA and an acetyl-CoA enzyme which undergoes 6 more rounds of beta oxidation to get the completely oxidized to yield the 7 more acetyl-CoA molecule. All the acetyl-CoA molecule produced into the beta oxidation of a single fatty acyl-CoA molecule get further oxidized in phase cycle to yield the NADH and FADH₂.

This means the 1 molecule of acetyl-CoA produces 3 NADH molecule, 1 FADH₂ molecule and 1 ATP or GTP. This means the 8 acetyl-CoA molecule is going to give you the 24 NADH molecule, 8 FADH₂ molecule and 8 ATP. Overall reaction for a permutoyl-CoA can be represented as follows. Permitoyl-CoA plus 7 CoA plus 7 FAD plus 7 NADH plus 7 water molecule will give you the 8 acetyl-CoA plus 7 FADH₂ plus 7 NADH plus 7 hydrogen molecule. This means if you started with the C16, it will actually go to leave 1 molecule of acetyl-CoA and actually go to small by the 2 atom.

Same is true by all these, right. And ultimately what will happen is that when the C14, it is actually going to produce the acetyl-CoA and it is also going to produce the remaining molecule is going to be the acetyl-CoA. This means the beta oxidation is actually going to generate a huge quantity of the liquid molecules are actually going to generate a huge quantity of energy, right. Remember that from the 1 acetyl-CoA, right, 1 acetyl-CoA you are actually going to generate 15 ATP molecule, right under the crate cycle. This means if I have 1 oxidation of 1 permethoyl-CoA, right, 1 permethoyl-CoA, it is actually going to generate the 8 ATP, 8 acetyl-CoA molecule. This means it is actually going to generate approximately 120 ATP molecules, right, from the beta oxidation.

Whereas it is also going to generate some or more amount of NADH and FADH₂ even from the beta oxidation step as well. Now, how you are going to do a regulation of fatty acid biosynthesis and catabolisms. So fatty acid regulation biosynthesis and catabolism is completely been regulated by the location of the fatty acids. So in the liver the fatty acyl-CoA has 2 major pathways. It can either transported to the mitochondria via the carnitine shuttle to get oxidized or it can be converted into the thioacylglycerol and phospholipid via the cytosolic enzyme.

The carnitine shuttle which is a 3 step process is the rate limiting step for the fatty acid oxidation and therefore it is an important point of regulation. Once the fatty acids are transported to the mitochondrial matrix, they are destined, designated for the beta oxidation. So remember that if the carnitine shuttle is not going to be working or if it is not functional, then the lipid molecules or fatty acid will not enter into the mitochondria for beta oxidation. Instead they will go for the cytosolic enzyme and they will be utilized for the synthesis of the phospholipids and the tricyclic drops. And these are the storage molecules or sometimes the phospholipids are going to be a part of plasma membrane.

So mannoly-CoA, the first intermediate of the fatty acid biosynthesis via the acetyl-CoA also regulate the fatty acid oscillation. When there is an ample amount of glucose supplied to the liver, fatty acid synthesis begins from the acetyl-CoA which produces the

mannoly-CoA that inhibits the carnitine acyl transferase 1. So this means if you have enough amount of glucose, so that means the glucose is enough to give you the energy and in that case what will happen is that the acetyl-CoA is actually going to be withdrawn from the crepe cycle. So this means if you have enough quantity from the glucose oxidation, there will be an access of acetyl-CoA. And this acetyl-CoA then would be working as a precursor for the fatty acid biosynthesis.

And this acetyl-CoA, the first molecule what it is going to produce is the mannoly-CoA. And mannoly-CoA is actually an inhibitor of the carnitine acyl transferase, the first enzyme which is in the carnitine shuttle, right. And if the first enzyme in the carnitine shuttle is not working, it is not going to attach the carnitine to the fatty acids and as a result, there will be no transport of fatty acid from the cytosol to the mitochondria and as a result, it is not going to go through to the beta oxidation. Instead, it is actually going to be utilized for the fatty acid biosynthesis. This means it is going to be utilized for the synthesis of triacyclic role or the synthesis of the other phospholipids.

So when the NADH and NAD plus ratio is very high, this means the cell is sufficient enough with the energy, okay. So you can imagine like that, okay. If it is a very high ratio, this means you have more amount of NADH and you have less amount of NAD plus that means the cell has sufficient energy, then it is indicating the enough energy for the cell to perform vital activities. Beta-hydroxycoenzyme dehydrogenase is also been inhibited. High concentration of acetyl-CoA inhibits the thiolase, right.

So these are the enzyme or which are actually functional within the for during the beta oxidations. During the time of vigorous muscle contraction, the stimulus exercise or fasting the consumption of ATP is increased which reduces the concentration of ATP and increases AMP that activates the AMPK, the AMP activated protein kinase. And AMPK phosphorylate various other targets enzymes such as stile-CoA carboxylase which catalyzes the mineral-CoA synthesis. This phosphorylation and thereby inhibition of coagulate carboxylase bring down the concentration of mineral-CoA relieving the inhibition of fatty acid acylcarnitine transporters into the mitochondria and allowing the degradation of the stored fat to undergo oxidation to regain supply of ATP from the fats.

So these are the things, right. If you have the high glucose molecules, you are actually going to produce the insulin and insulin is actually going to participate into the fatty acid biosynthesis and regulation. So when the blood glucose level is high, the insulin dependent protein phosphatases dephosphorylate the acetyl-CoA carboxylase thereby activating it and ACC starts sizing the mineral-CoA which inhibits the carnitine acyl transferase 1 and thereby preventing the entry of fatty acid coenzyme, fatty acid CoA, fatty acyl CoA into the mitochondria. This means once you have the high blood glucose

level, it is actually going to induce the production of insulin and once there will be an induction of insulin, insulin will actually go and bind to the insulin receptor and that in turn is actually going to produce the large quantity of phosphatases. And once the large quantity of phosphatases is produced, it will actually go to dephosphorylate the acetyl CoA carboxylase. So acetyl CoA carboxylase is inactive when it is phosphorylated and it is active when it is dephosphorylated or native form.

So once the phosphatases are produced, they are actually going to have the active ACC and what is the job of the active ACC is that it is actually going to take up the acetyl-CoA from the crepe cycle and it is actually going to produce the manolil-Ka and manolil-CoA which is actually going to form the fatty acid and manolil-CoA is actually a very, very potent inhibitor of the carnitine acyl transferase 1. So this means it is actually going to destroy the carnitine shuttle and once it is destroying the carnitine shuttle, it is actually going to destroy the transport of fatty acid into the mitochondria and if it is destroying the entry of the fatty acid into the mitochondria, it is actually going to abolish the beta oxidation of the fatty acids. And as a result, it is actually going to promote more the synthesis rather than the degradation of the fatty acids. So that is why it is always been recommended that if you want to reduce the amount of fat into your body, you always should ensure that there is no enough glucose present. This means what it means is that it is not the fat which actually increases the fat level, it is the glucose which actually increases the fat level because if you have a high quantity of glucose within the blood, it is actually going to promote the synthesis of fat rather than the fat burnout, right.

And that is why it is important that we should have the less amount of glucose into the blood. So when the blood glucose level drops, the glucagon release activates the pKa which phosphorylates and inactivates the ACC. The concentration of the melanin coate drops which leaves the inhibited entry of fatty acid into the mitochondria and replenishes the beta oxidation. So this is all about the catabolic reactions what we have just discussed, right. And what we have discussed, we have discussed about the catabolic reaction of the glucose and the lipids or the fatty acids and within the glucose, we have discussed about the glycolysis and we also discussed about the crepe cycle.

Whereas in the case of lipid molecules, we have discussed about beta oxidation and how the beta oxidation is producing the acetyl CoA and then if this acetyl CoA is entering into the crepe cycle and that it is now it is actually going to produce the large quantity of energy in the form of ATP. So now once you have generated the large quantity of ATP, this ATP is actually going to be utilized into the anabolic reactions. Anabolic reaction means the biosynthetic reactions and anabolic reactions are required for the growth of the organism or growth of the person, right. Because if you want to grow for example, if you want to grow from 1 mm to 1 centimeter, right. This means you actually require the

enough quantity of material, right.

So that you can actually be able to for example, if you want to increase the length, right. So you also have to synthesize the muscles, right. And muscles is nothing but made up of the different types of protein molecules, lipid molecules, right. So you also require the synthesis of protein and lipids and you also require the nucleic acid, right.

So if you want to do a synthesis, you also require the energy. So energy you have already produced, right. And this all energy is the endogenous energy, it is not the exogenous energy, right. And I am sure you all very much aware of what is mean by the endogenous energy, what is mean by the exogenous energy, right. For example, if you take a carbon molecule, if you take a carbon, right, and if you this, if you burn this carbon, it is actually going to give you the energy, right. This is exogenous energy, because it is always present outside, right.

That is how you do actually, you take the carbon, you burn it, and that is how you keep it into corner of your room and that actually keep the room warm actually. But this is not going to remain continuous, right. Whereas in these cases, you are actually producing the energy by running the different types of metabolic reactions within your body, right. So these are the endogenous energy. And that energy is actually going to be utilized for forming the bonds between the constituents.

For example, in this case, if you want to make the protein, you always have to make the bond between the amino acids. If you want the lipid synthesis, you always have to make the bond between the fatty acid and glycerol. Similarly, for the nucleic acid, you always have to make the bond between the nucleotides molecule. And that is how it is actually going to synthesize genome for the new cell, it is going to synthesize the plasma membrane for the new cell, it will also require the synthesis of the protein for the new cell. And once you have all these raw materials, the cell is actually going to enlarge in size and that is how it will actually going to help into the growth of the organism.

So let us discuss about the anabolic reactions. And what we are going to discuss, we are going to discuss about the biosynthesis of the amino acids, where all this energy is going to be utilized. So amino acid biosynthesis. Amino acids are categorized into the two different categories, essential and non-essential amino acid based on the biosynthesis. Thus, the amino acids which are actually you which you can be able to synthesize in your body by the raw material are called as the non-essential amino acids. Because these are the amino acids which you can be able to synthesize from the raw material.

Whereas the amino acids for which either you do not have the biosynthetic pathway or

you cannot synthesize from the raw material, because you do not have the requisite biosynthetic pathway, they are called as essential amino acids. So for the plant, plant can be able to synthesize all the amino acids. So it is actually all the amino acids are non-essential amino acids for the plant because plants can easily take the carbon dioxide, water and the other metabolites and it can be able to synthesize all amino acids. Because the plant has the biosynthetic pathway for all amino acids.

Whereas the animals are dependent on the plant to provide the amino acids. So these are the amino acids which are essential. So these are the amino acids for which there is no pathway present in the animals. This means these are the amino acids which it has to take from the plant. So the plant has to provide these and how the plant provides these? Plants are actually giving you different types of raw material, for example pulses. So if you take the pulse, pulse is actually going to be digested into the digestive system and it is actually going to release the different types of amino acids.

For example, if you take the rice, rice is also going to produce some amount of amino acids which are funding into the essential amino acids. And then you have the non-essential amino acids where you actually have the biosynthetic pathway. So you have the pathway and you just require the raw material. So you actually require the ammonia, you require the carbon dioxide, you require all those kind of things and then you can be able to produce these amino acids or sometimes some of these amino acids are also being derived from the essential amino acids.

So that is also you can actually be able to have the biosynthetic pathway. So these are the 10 amino acids and these are also 10 amino acids. So biosynthesis of the amino acids. Principally, all amino acids are derived either from the glycolysis or citric acid cycle or pentose phosphate intermediates. These derivatives provide the carbon skeleton from the amino acids whereas amino group or the nitrogen in the same is provided either by the glutamine or the glutamate. Not all amino acids are synthesized by the organisms which they need from the outer environment either in the form of protein or from the dietary food.

These amino acids which they cannot synthesize by the organisms are called as essential amino acid and the rest are called as non-essential amino acid. The most important reaction that take part in almost all the biosynthetic pathway of different amino acids are reductive amination of the alpha-keto acids or the transaminations reactions or require a coenzyme PLP that is a pyridoxyl phosphate. So this is an overview of the amino acid synthesis where you actually have the different types of amino acids derived from the carbohydrate metabolism. This is the glycolysis what you see here is from here to here right this is the glycolysis and then from the pyruvate this is the crepe cycle okay. And

what you see here is that from the first for example from the glucose you can actually be able to have the ribulosis phosphate and from that you can actually be able to have the synthesis of histidine.

Similarly from the 3-phosphoglycerate you can have the synthesis of serine and once you have synthesized the serine that serine can be converted into the glycine and the cysteine. Similarly 3-phosphoglycerate can enter into the pentose phosphate pathway and then pentose phosphate pathway is going to generate the Rheicose 4 phosphate and that along with the phosphoenolpyruvate can give you all the aromatic amino acid like tryptophan, phenylalanine and tyrosine and then we also have the other amino acids like aspartate, asparagine, methionine, theanine all that from the crepe cycle like the oxaloacids right. Similarly from the alpha-glutarH you can actually be able to produce the glutamate by the different types of metabolic reactions and once you produce the glutamate you can convert that into the glutamine, proline and arginine. Similarly from the pyruvate you can be able to produce the alanine, valine, leucine and isoleucine. So based on these kind of scheme the amino acid biosynthesis can be decided into the different families.

So you can have the glutamate family, you can have the pyruvate family, you can have aspartate family, serine family, aromatic amino acids and you can also have the histidine families. So when you talk about the glutamate family you only what you need is you require the synthesis of the glutamate. Once you synthesize the glutamate you can be able to synthesize the glutamate, glutamine, arginine and proline. Similarly once you have the pyruvate, from the pyruvate you can be able to generate the valine, alanine, leucine and isoleucine. From the aspartate you can be able to synthesize all these, from the serine you can be able to synthesize all these.

From the aromatic amino acids you can be able to have the tipto-phan, phenaniline and thiocine. So let us start first with the glutamate family. So biosynthesis of the glutamate and glutamine. So from the alpha-ketoglutarate which is present in the crepe cycle, you can be able to synthesize the glutamate and once you synthesize the glutamate you can be able to convert that glutamate into glutamine, proline or the arginine.

So biosynthesis of the glutamate and glutamine. So glutamine synthesis is an important mechanism of ammonia assimilation, transportation in different cells and secretion therefore after. So free ammonium ion, free ammonia is toxic for the cell which is converted into glutamine for the transportation. In bacteria and plant, the glutamate is derived from the glutamine catalyzes, from the glutamine catalyzes by enzyme known as GOGCAT or the glutamate oxo-glutarate aminocharsinase. Here the glutamine act as a nitrogen donor and alpha-ketoglutarate undergoes the reductive etiaminations. So what

you have is to have the alpha-ketoglutarate, glutamine, NADPH, ATP, remember this is NADPH not the NADPH and it is actually going to produce the two molecules of glutamate plus NADPH.

So you are actually going to utilize not only the ATP but also in the form of reducing equivalents. Animals do not have the glutamate synthesis, therefore they maintain the high level of glutamate by transamination of the alpha-glutarate while the amino catabolism. Glutamate can also be formed by the glutamate dehydrogenase in the single step reaction given below. The reaction takes place in the mitochondria, the reaction cannot distinguish between NADH and NADPH.

So alpha-ketoglutarate with ammonia, NADPH give you the glutamate and NAD+. Then we have the biosynthesis of the serine, glycine and cysteine. So from the 3-phosphoglycerate, so 3-phosphoglycerate you are going to get from the glycolysis and from the glycolysis then the 3-phosphoglycerate can be produced into the serine and serine can be converted into glycine and cysteine. So biosynthesis of serine, glycine and cysteine. So serine is derived from the oxidation of 3-phosphoglycerate by the phosphoglycerate dehydrogenase in the presence of NAD+, to produce the 3-phosphoglycyl pyruvate and glutamine, since this transfer its amino group to the above synthesized product to yield the 3-phosphoserine followed by the hydrolysis of phosphate group by the enzyme called as phosphoserine phosphatase to yield the serine. The pathway for serine and glycine are almost the same except for the synthesis of glycine after the removal of carbon atoms from the serine by an enzyme called serine hydroxymethyl transferase or the SHMT.

In the above reactions, the beta carbon of the serine is accepted by the tetrahydrofolate in the presence of PLP. In plant and bacteria, cysteine is derived from the serine and for which the sulphur is obtained from the environmental sulphates. First, an acetyl group is attached to serine from the acetyl-OA to form the O-acetylserine. This reaction is performed by the enzyme serine acetyltransferase. The reduced sulphur is then incorporated into our product by enzyme called O-acetylserine thiolase to yield the cysteine.

In mammals, the process is quite different. The carbon skeleton and the sulphur for cysteine biosynthesis is given by the two different amino acid that is the serine and the methionine respectively. So, these are the reactions what is being shown here. You start with the 3-phosphoglycerate, it gets converted into 3-phosphohydroxypyruvate and then the glutamate is going to give you the amino group and that is how it is actually going to form the 3-phosphoglycerine and thioeposocerine is actually going to be hydrolyzed and it is actually going to give you the serine. And once the serine is being produced with the

help of the enzyme called serine hydroxymethyltransferase is going to be get converted into the glycine. And remember that tetrahydrofolate or N-5-antennal-metallentetrahydrofolate is actually a shuttle between the nucleic acid and the protein molecules.

That is why this enzyme is actually can be shuttle the carbon pool between the protein metabolism as well as the nucleic acid metabolism. And then we have the biosynthesis of the aspartate family amino acids. So oxalacetate you have transaminations to form the aspartate and on the aspartate you can have the amidation to produce the arginine, asparagine from the aspartate you can have the methionine, threonine and lysine. Then we have the biosynthesis of the pyruvate family amino acids. So pyruvate family, pyruvate after the transamination can produce the alanine and the pyruvate can produce the valine, isoleucine and leucine.

So and then we have the biosynthesis of the aspartate and alanine. So carbon skeleton for aspartate and alanine is derived from the oxalacetate and pyruvate respectively whereas amino group is provided by the glutamine for both the amino acids. In the above reaction, the alpha-ketoglutarate is formed as the byproduct along with the alanine and aspartate as amino acid. This is an example of the transamination reaction and it is catalyzed by the amino transferase in the presence of coenzyme PLB. So you have this kind of transamination reactions where one side you have the oxaloacetate and glutamate and this side you are going to generate the aspartate and the alpha-ketoglutarate.

So it can actually go in both the directions. So depending upon whether you require the aspartate or whether you require the glutamate, it can actually be able to convert the enzymes, convert the amino acid into each other. Same is true for the pyruvate and alanine also. And then we have the biosynthesis of the proline and the arginine. So these are the reactions what you are going to have for the biosynthesis of the proline and arginine.

So it starts with the glutamate. Glutamate is actually going to be having the energy of the ATP and it is actually going to form the gamma-glutamyl phosphate and there will be a phosphorylation reaction. So it is going to be catalyzed by an enzyme called glutamate kinase. And once you generate the gamma-glutamyl phosphate, then gamma-glutamyl phosphate is going to be reduced by an enzyme called gamma-glutamyl phosphate reductase and that is going to form the glutamate semi-aldehyde. And then this semi-aldehyde is actually going to be get converted into the proline-5-carboxylase carboxylate and this 5-phosphoryl carboxylase is further going to be reduced by enzyme called proline-carboxylase reductase and ultimately it is going to form the proline.

In animals, the arginine is produced from the glutamate in the urea cycle. Similarly arginine is derivative of the ornithine which can also be produced from the glutamate gamma semi aldehyde by transamination reaction. But cyclization of gamma semi aldehyde interdict the enough supply of the same to synthesize the ornithine. In the case of bacteria, there is a de novo pathway altogether for the formation of ornithine and therefore arginine. So biosynthesis of the aromatic amino acids. So from the phosphoenolpyruvate, it is actually going to combine with the archaeosis phosphate from the pentose phosphate pathway and that is how it is actually going to form the tryptophan tyrosine and phenylalanine and tryptophan can further be converted into tyrosine if required.

So from the chorismate, it is actually going to form the anthraenoallene and then from here it is actually going to form the enoyl road carboxyphenyl amino carboxyribulose phosphate and from here it is actually going to form the indole-3-glycerol phosphate and ultimately it is going to be get converted into the tryptophan. So once it is going to form a tryptophan, it can be get converted into the tyrosine. So enzyme tryptophan synthase which performed the last reaction in the conversion from the chorismate to tryptophan has two subunit alpha 2 beta 2 that perform the two different parts of the whole reaction. Indole-3-glycerol phosphate alpha subunit is going to form the indole plus geseltii plus 3 phosphate and indole plus terrine is actually going to form the tryptophan and there will be a bowl of water. So the bowl formed in the first part of the reaction is moved to the channel from the alpha subunit, the beta subunit activate active site where it undergoes condensation with shift base intermediate such as the PLP and serine and in any one the tyrosine can be formed by the hydroxylation phenylalanine at C4 position by enzyme called phenyl hydroxylase.

So this is the tyrosine and when you have the phenyl hydroxylase, it is actually going to form the, so this is the phenylalanine actually and when it has the phenylalanine hydroxylase, there will be hydroxylation reaction. So it is going to have the hydroxylation on this side and it is actually going to form the tyrosine. Then we have the biosynthesis of phenylalanine and tyrosine in plant and bacteria. So in the plants and bacteria, the phenylalanine and tyrosine is derived from the chorismate where prefinite is a common intermediate and then the pathway diverges to the two branches, one forming the tyrosine from the 4-hydroxyl pyruvate and the other forming the phenylalanine from the phenyl pyruvate.

The final reaction is the transamination that involves the transfer of the amino group from the glutamine. So this is the pathway what is being shown from the chorismate, it is going to form the p-phenate and at this stage, it is going to bifurcate into the two pathway and one side it is going to form the 4-hydroxyl pyruvate, other side it is going to

form the phenyl pyruvate and then it is actually going to have the transamination reaction and that is how it is going on this side it is going to form the tyrosine whereas in this side it is actually going to form the phenylalanine. So regulation of the amino acid biosynthesis, so you can have the regulation of the different types of biosynthetic pathways and it all depends on the availability of the different types of metabolites and that is how they are actually going either going to upregulate or down regulate the different enzyme activities. For example, in this case the glutamine synthase, so amino acid biosynthesis is aerostically regulated. The end product of the pathway generally regulate the enzyme that catalyze the initial step of the pathway.

Along with the acrylic modulation, feedback inhibition is also been seen to regulate the amino acid biosynthesis. The glutamine synthase is an important enzyme that participate in almost all the reaction with all of the amino acid biosynthetic reaction. Therefore, this enzyme is inhibited by the various other molecules such as AMP, CTP, glycine, allene, etc. The other mechanism seen are the sequential feedback mechanism which are more profound in the aromatic amino acid biosynthetic pathway. And this mechanism, the amino acid phenylalanine, tyrosine, tryptophan, sequentially inhibit the 3-isoline of the enzyme THV and that is how it is actually going to inhibit the synthesis of the aromatic amino acids. So this is all about the amino acid biosynthesis and what we have discussed, we have discussed about the catabolic reactions and we have also discussed about the anabolic reactions.

And so the purpose of the metabolic reaction is to produce the energy whereas the purpose of the anabolic reaction is to utilize that energy for the synthesis of the different biomolecules which they require for the synthesis of the different types of biomolecules such as proteins, lipids and the genome. And all this is required for the synthesis of the new cell so that they can be able to grow from the unicellular organisms to multicellular organisms or they can actually be able to produce the more number of cells so that they can be able to increase their number. So with this, I would like to conclude my lecture here. In our subsequent lecture, we are going to see more aspects of the biological system. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 02
Basics of Biological system (Part 2)
Lecture-08 Cell Cycle and Control (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati and what we were discussing, we were discussing about the basics of the biological system and in this context so far what we have discussed, we have discussed about the cellular structures. So, initially we discuss about the prokaryotic structures followed by the eukaryotic structures and when we were discussing about the eukaryotic structures, we have also discussed about the different organelles what are present in the eukaryotic cells and how you can be able to separate them with the help of different types of fractionation techniques. We have discussed about the density gradient identification and as well as the differential centrifugations and in the previous lectures, we have also discussed about the cellular metabolism, so we have discussed about the anabolic reactions and we also discuss about the catabolic reactions. So, within the catabolic reaction, we discuss about the carbohydrate metabolisms and lipid metabolisms whereas within the anabolic reactions, we discuss about the synthesis of the different types of amino acids and so on. So, in today's lecture, what we are going to do is we are going to discuss about the how the cell is utilizing this energy for different types of activities and one of such different activity which is very crucial for the cell is the cell division and the cell cycle.

So, when we talk about the energy, the energy what is present in the biological system is going to be utilized for many types of applications, this energy is going to be utilized for the synthesis of the anabolic reaction and the end product of the anabolic reaction is the synthesis of biosynthesis of product, biological molecules and such as protein and nucleic acid. Most of these things are required for the other kinds of activities. Now, energy is also being utilized for the growth of the organisms. So, when you take a nutrition, you are going to produce the energy and that energy is going to be utilized for anabolic reaction, running the anabolic reactions, so that you can be able to synthesize the different types of biomolecules and that energy is also going to be used for the growth of the organisms.

And when you are doing the biosynthesis of the different types of biological molecules, that also is going to be utilized for the growth of the organisms. Because when you are going to increase the number of cells or when you are actually going to enlarge the growth of the cell, you require the synthesis of the plasma membrane, you are going to

require the synthesis of the nuclear content and then you also require the synthesis of the different types of molecules. So, when we talk about the prokaryotic cell, in the prokaryotic cell, let us take an example of the bacteria for example. So, if you take a bacteria, what happened is that the bacteria is having the chromosome right or the nuclear content right. So, what happened is that when it actually requiring the nutrition from outside, so it is taking the nutrition right, it is taking up the nutrition and then enlarging into the site.

So, it is synthesizing the lipid molecule, it is synthesizing the protein molecules, it is synthesizing the nucleic acids and all that and as a result what is happening is that it is actually going to increase the size right, it is actually going to increase the cell size of the cell and the nuclear DNA also right. When it goes to beyond a certain limit okay, then what happened is that it is actually going to make another copy of the DNA and as a result what will happen is that it is going to have the two copy of the DNA right and once it has the two copy of the DNA, it is actually going to be divided from the center right and that is how it is actually going to have the two different bacteria right and with its own nuclear DNA, its own chromosome right and that is how the bacteria is going to grow from one bacteria to another bacteria and this kind of division is called as binary division right. And binary division could be equal division or it could be unequal division, apart from the binary you can also have other kinds of divisions such as fragmentations and other things of that right but binary division is the most popular division what is happening into the prokaryotic system. Now, here when you are doing the binary division, you are simply cutting the bacteria from the middle and so that both the cells will have the some amount of cytosol and some amount of the nuclear content and ultimately these two are also going to take up the nutrition and then they also will grow right. But since they do not have the organelles, they do not have the membrane bound organelles, it is easy because you do not have to distribute the organelles also.

For example, if it is a eukaryotic cell, you also have to ensure that when you are going to do the cell division or when you are going to divide the cell and produce the two cells, you should have the equal distribution or nearly equal distribution of the mitochondria, you also should have the distribution of the endoplasmic gut equilibrium, you should also have the distribution of the Golgi bodies and so on. So that the both the cells should be sufficient enough or independent in terms of the you know running its metabolisms. And that is very difficult and that requires a precise mechanism so that you can be able to you know do the cell division and that is why the cell division in the eukaryotic cell is well planned and well organized in such a way that you are actually going to have a sequence of events so that you can be able to perform the cell division. And these sequence of events are called cell cycles. So every eukaryotic cell is actually going to go through with these cell cycle stages and then only it actually can be able to you know

increase its number.

So it can actually be able to divide and give you the more numbers. So the cell cycle when you talk about the cell cycle, the cell cycle has the precise different phases, you have the interface, you have the G₁, S, G₂ and M phases. So the eukaryotic cell undergoes the precise cell cycle and division to produce the two daughter cells. Cell cycle is the series of tightly regulated events leading to the division and duplication. It is a vital process used to single cell called fertilizer egg is developed into the full organisms.

Several division is the crucial event underlining the regeneration and repair in the tissue, liver and heart. In eukaryotic cell, the parent cell is divided simply by the division in the two halves through the process of binary fission. Whereas in the case of eukaryotic cell, the cell cycle has three important phases, the interface, the mitotic phase and the cytokinesis. So in the interface, you are going to have the synthesis of the genomic content and the tetraploid. Whereas you can in the mitotic phase, you are going to have the division of DNA into the two halves and the cytokinesis, it is actually going to have the division of the cell.

This means in a cell cycle, you are going to have the interface. So this interface is actually going to prepare the cell for the cell division. And you can also have the S phase where you are actually going to have the synthesis of the DNA and the G₁ and G₂ phase are actually the and then you also going to have the mitotic phase which is actually going to divide the DNA into the daughter cells and then it also going to be followed by the cytokinesis which actually going to divide the cell into the two cells. So let us start discussing about these different phases. So first is the interface.

So interface is preparatory phase required to perform the requisite steps. These are the series of events in the nucleus and as well as in the cytosol of the daughter cell to enable it to enter into the division phase. Division phase, this phase has several phases, there are as follows. So within the interface, you are going to have the G₁ phase. So it is also known as the growth phase, it starts from the end of the mitosis and until the beginning of the S phase.

So this is the G₁ phase. It starts from the mitotic and it goes up to the end of the beginning of the S phase. During this phase, the cellular proteins, enzymes are synthesized. So during this phase, you are actually going to have the synthesis of the different types of biological machinery, right, you require the DNA polymerase, you require the actin protein, you require the plasma membrane, all those kinds of things are actually going to be synthesized within the G phase. So G₁ phase is actually where there

will be a synthesis of cytosol so that it is actually going to help for the cell to enter into the S phase.

Most of these enzymes are required for DNA synthesis in the S phase. Duration of the G1 phase depends on the cell type within the organism. G1 phase is under the control of the p53 gene products, right. So you know the p53 is a transcription factor which actually regulates the cell cycle and the cell division. So the G1 phase is under the control of p53.

And G1 phase, the length of the G1 phase depends on the availability of the different raw material what is required for the synthesis of the enzyme. It also be regulated or it is also being linked to the genetics of that particular cell. So if the cell is very slow growing cell, then the G1 phase probably could be very large because it requires a large quantity of the enzymes and other kind of things. Whereas the fast growing cells, the G1 phase could be short because it has to divide very rapidly. And in that case, the G1 and G2 phase could be very, very small, because it actually is full of nutrition.

So it actually can prepare the G1 and G2 phase very quickly. Then we enter into the S phase. So the S phase is the one cells grow and all factor nucleotides available. It starts the DNA synthesis during the S phase. At the end of this all chromosome present in the nucleus is replicated and the DNA content get doubled.

No change in ploidy. The synthesis of DNA occurs very fast to avoid the exposure of newly synthesized DNA to the mutation. So in the S phase, you are actually going to have the DNA synthesis. DNA synthesis from the pre-existing copy. This means if you have the parent DNA, if you have a parent DNA, that parent DNA will enter and will give you the two DNA strands.

One is parent DNA. One is the original DNA or the parent DNA and the other one is going to be the daughter DNA. Right. And you know that the DNA is, the application is semi-conserved in nature. So the parent DNA is also going to have the new copy and the daughter DNA is also going to be the new copy. This means both of these DNA strands are going to be identical in nature, which means the different DNA, it will be going to have one strand from the parent DNA and the other strand from the newly synthesized DNA.

Whereas in the daughter cell also you are going to have the one strand from the parent DNA and the other strand is going to be from the newly synthesized DNA. And that is how the DNA synthesis is going to occur during the S phase. And at the end of the S phase, the amount of DNA is going to be the 2x. It is not going to be in change the

priority of the cell. It is actually just increasing the DNA content of that particular cell.

So earlier when it was in the G₁ phase, it is actually going to have the X amount of DNA. Whereas in the S phase, at the end of the S phase, it is actually going to have the 2x amount of DNA. Then it enters into the G₂ phase. So G₂ phase is also the preparatory phase, the growth phase between the DNA synthesis and the mitosis. During this phase, the cells grow and synthesize the protein and cellular machinery required for the mitosis and the cytokinesis.

This means it is actually going to produce large different types of the cyclin proteins and other kinds of protein which are spindle proteins, tubulins and all those kind of things so that they will help into the mitosis and as well as during the cytokinesis. Apart from these two phases, we also have a phase which is called as G₀ phase. So G₀ is actually a non-dividing phase. So it is a non-dividing phase. So after the G₁ phase, the quiescent, quiescent means the cell which are non-dividing cells, senescence which means the cell which are very sick or they do not want to divide and non-proliferating multicellular eukaryotic cell enters into the G₀ phase because they do not want to divide.

For example, the neural cells, right? And so the cells of the brain, the first cells of the spinal cord, the cells of the neural system, they do not divide because they are terminally being programmed like that. So they will always remain into the G₀ phase and they will never enter into any of these cell cycle stages. So cells remain in this phase for long period of or the indefinite period as in the case of the neuron cells. It is also common in fully differentiated cells. The fast growing cells never enter into the G₀ phase and hence it is not a regular cell cycle phase and the cells are undergoing into the specific conditions they will undergo into the G₀ phase.

For example, the RBC, RBC is under the G₀ phase because it does not decide, because RBC does not have the nucleus. So it has actually a specific case where a particular type of cell does not have a nucleus and that is why it cannot divide. So it will actually enter into the G₀ phase. Apart from that, the neural cells are also non-dividing cells. So they still have the nucleus, but since they do not get any kind of stimulus from the external so that they can actually be able to decide to divide actually.

So after the G₂ phase, you are actually going to have the mitotic phase and then you also going to have the cytokinesis. So mitotic phase, whether it is, so mitotic phase can be divided into two different types of phases. It can be mitosis or the meiosis. So mitosis or the M phase after the G₂ phase, the cell enters into the mitosis or the M phase to divide the DNA equally between the two daughter cells. Each mitosis has the four distinct states to precisely divide the DNA content of the cell.

So the purpose of the mitotic phase is that it wants to divide the DNA precisely between the two daughter cells. So it actually can divide in such a way that the both of these cells are actually going to get the equal amount of DNA. So the mitotic phase is also being divided into the four different phases. First phase is called prophase. So during this phase, the nuclear membrane is dissolved and the chromatin condense into the chromosome.

The nucleus nucleolus in the nucleus disappear. In the beginning, each cell has two centromere, each cell has one centromere, which replicates along with the DNA to give rise a pair of centrosome to coordinate downstream events. Each centrosome has microtubule to form the spindle and assist in the distribution of the nuclear content during the mitosis. Centrioles are considered to organize a microtubule assembly, but they are not essential. So this is what you are going to see here is the different types of phases within the during the mitosis.

So you what you see here is the mitosis. This is the prophase, then metaphase, then anaphase and telophase. During the telophase, both the chromosomes, both the DNA strands are going to be separated and that is how it is, then you are going to have the cytokinesis. So then it is actually going to divide and get you. And what you see here is the, you know, the different types of cells. And this is a typical pattern what you are going to see because some of the cells are under the M phase, some cells are under the prophase, metaphase, anaphase and all that.

So this anyway, we are going to discuss in detail when we are going to talk about the experimental top setup. So after the prophase, you are going to have the metaphase. So in the metaphase, so once you have, you know, loosened nucleus, so during the prophase, what you have done is your first thing what you have done is you have removed the nuclear membrane, so that the DNA what is been synthesized and the two copies of the DNA what is present inside the nuclear is now free for distribution, right. And then you are actually going to, you know, pick up the DNA, right. So you're going to divide the chromosome into the two parts.

So in this phase, the two centromere start pulling the chromosome using the attached centromere towards the end of the cell, which means what happened is that you are actually first going to dissolve the nuclear membrane, so that the magnetic content what is present inside this is actually going to be freely be accessible by the cellular machinery. So what happened, this is what you're going to happen during the prophase. In the beta phase, what will happen is that the chromosome is actually going to be attached and it starts pulling the chromosomes, right. So you have two copies of the

chromosome of the same chromosome and it is going to be start pulling on to the end of the cell, right.

So imagine that you have this as a cell, right. So one end will enter into this side, the other DNA will enter into this side. And that will happen with not with the one chromosome, but it will happen with all the chromosomes, right, all the chromosomes will be going to fold into the two poles of the cell. So as a result, what will happen is that the chromosomes are aligned along the metaphase plate or the equatorial plane. Since the pulling power of both the centrosome is almost equal, it eventually arranged the chromosome on the metaphase plate. So this is what's going to happen, it is actually going to arrange all the things on to the metaphase plate.

The alignment of the chromosome along with the metaphase plate is crucial event to decide the entry of cell into the another phase which is called as the anaphase. The signal required for this control is created by the mitotic spindle checkpoints and all these checkpoints are being controlled by the cycle independent cell cycle proteins. Then we have the anaphase. So the protein attached to each chromatids are cleaved and the sister chromatids are separated as the daughter chromosome. So chromosomes linked on the metaphase plates are pulled by the microtubule and move towards their respective centrosome.

Although the exact mechanism of generating the force required for the centrosome movement is unknown, but it is suggested that the rapid assembly and breakdown of microtubule may provide the force for this movement. At the end of this says that chromosomes are being prepared for the distribution between the two different types of cells. And then we have the telophase. So in the telophase, in this phase, the daughter chromosomes moved and attached to the opposite end of the cell, right.

So this is what this is the telophase. The nuclear membrane forms around each set of the separated chromosome daughter chromosomes and the nucleothorolus reappears. In this event, the several processes during the prophase are reversed to give the two daughter nuclei. So this is what going to happen. After this, you are going to have the cytokinesis.

Cytokinesis means the division of the cellular content. At the end of the telophase, the mitosis is over, but the cell division requires the distribution of the cellular content equally between the daughter cell. In the animal cell, a cleavage photo is formed between the along the metaphase plate and divide an individual nuclei as the separate cell. During this process, it is ensured that besides nuclei, all other cellular organelles should be distributed equally between the daughter cell. Whereas in the plant cell, the

cell plate is formed and divide the cellular content between the daughter cells. So cytokinesis is a very, very important step or very, very important events which actually going to divide the cellular content, such as mitochondria, endoplasmic reticulum, Golgi bodies, centrosomes, paroxesomes and all those kinds of things.

So that both of the components are actually going to be equally competent or equally independent so that they can be able to run their own metabolisms. Now, so these are the four different stages what are present into the cell cycle. You have the we started with the G1, then the cell are actually going to prepare for the cell division, cell is going to be prepared for the DNA synthesis during the G1 phase. So there will be a synthesis of DNA polymerase, nucleotides and all that kind of. Once that is ready, the cell will enter into the S phase and within the S phase, it is actually going to do the replication.

And once the cell has replication, it is actually going to receive the X amount of DNA from the G1 phase. And that is actually going to be get converted into 2X amount of DNA during the S phase. That 2X will enter and go to the G2 phase and then it will enter into the mitosis. So during the G2 phase, it is actually going to synthesize the machinery what is required for performing the M phase or mitosis. And in the mitosis, then it is actually going to divide the nucleus and it was actually going to distribute between the mother cell and as well as the daughter cell.

It means the amount the DNA content is again going to be 1X at the end of the mitosis. And then there will be a cytokinesis so that it is actually going to produce the daughter cell and it is also going to have the mother cell which will enter into the cell cycle. So if I want to study this particular type of phenomena during the using the some of the analytical tools, what we require is we require a machinery so that it can be able to differentiate the cellular content. So you know that from this to this, the cell will actually grow.

So it actually going to change the size also. And it is also going to change the DNA content. So we require a machinery so which can actually be able to monitor the size and it also can monitor the DNA content of the cell. And that is how you can be able to, you know, identify the different types of cells what are present in the G1 phase, S phase, G2 phase and M phase. And within the M phase also you can actually be able to decide which cells are under the prophase, metaphase, anaphase or telophase. So one of the classical technique what you can actually be able to do for studying the size and the DNA content is called as the flow cytometry.

And the flow cytometry is a very, very robust tool to study the different types of cellular properties. So before getting into the detail of how you can be able to study the cell cycle

using the flow cytometry and how you can be what are the different types of protocol what you have to follow, I would like to show you some of the basics of the flow cytometry. So flow cytometry is a very robust tool which actually studies the different types of activities. So it actually can measure the density, size, it can also be able to tell you the sector what are present in the cell surface and it also can be able to, you know, differentiate the cell based on the metabolic reaction. How it actually happens? It happens because it has a cell analyzer and cell analyzer is actually streamlining all the cells into a small chamber and once the cell will exit out of this chamber, it is actually going to be illuminated by a laser.

So once it is going to be illuminated by the laser based on the size or based on the all these properties like the cell surface receptors or the metabolic reaction, it is actually going to give to the signal to the different types of detectors. So it can actually give you the size to the detector for measuring the size, it can actually give the signal to the detector for measuring the density or it actually can give the signal to the different types of fluorescence signal. And accordingly, you can be able to know which cell has the B cell receptors or which cell has a T cell receptor and so on. And based on these kind of informations, you can actually ask the machine and it can actually be able to collect these cells which are would be having the desirable features. So to do this job, we have a very detailed instrumentation part where you are actually going to have the different types of components and the cells are actually going to flow into a chamber or into a flow and once they reaches into the center and it is actually going to be illuminated by a laser.

And once it is going to illuminate by the laser, the cell is actually going to reflect the signal and it is actually going to be captured by the different types of the photomultipliers and these photomultipliers can give you the signal for the different types of properties. So it can be give you the information about the size, density and all other kinds of things. So it actually has the 3 main components, you can have the flow system which is called fluidics, you can all have the optical system and you can also have the electronic system which means whatever the signal you are going to get, it is actually going to process the signal and it will give you the readable signal. So as far as the flow system is concerned, it is going to have the flow cell and that flow cell is going to have the central core which is sample is going to injected and then also going to have the outer sheet membrane and it is going to have that and because of the hydrodynamic focusing, it is actually going to focus the cell in such a way that they will be actually going to travel in a single stream and that is how the single cell is actually going to be illuminated by the lasers present at the end of the tube. As far as the optical system is concerned, the light source used in the flow cytometer could be either the laser beam or it can also have the arc beam.

You can also have the organ lasers which are actually going to give you the 488 nanometers wavelength. Although in modern flow cytometers can actually be able to have the different types of lasers, can have the 488 lasers, can have 280 lasers and all that. So this is actually making the things more robust so that you can be able to collect more information about the cell. Then we also have the different types of detectors. So you can have the detectors for the forward scattering, you can also have the detectors for the cell size, reflective index and so on.

Then you also have the detectors for the side scatters and then you also have the intensity and whatever the intensity you are going to get, this is going to be the signal for the cytosolic content of the structure. And then you also have the electronic system. So electronic setup is going to be convert the information or the photons to the photoelectrons and these measure the amplitude area and the width of the photoelectron pulse and it amplifies the pulse either linearly or the logarithmally and then it supplies the amplified pulse. Then we also have the different types of plots what you are going to show into the flow cytometry. You can have the single color histogram, you can have the two color dot plot, you can have two color contour plot or you can also have the density plot.

And in the data acquisition when you are actually going to prepare a sample for the flow cytometry, what you are going to do is you are first going to do a you know trial run or test run actually right. And using the test run you can be able to set up the voltage, you can be able to set up the gain in such a way that it should actually you know your untreated sample should be in the center of the graph okay. And so that you can be able to monitor the moment in the left or the right directions right. Apart from that you can also do the gating right and gating is actually nothing but the selection of a subset okay.

Selection of a subset. So selection of only a certain population of cell for analysis on a plot right and it allows the ability to look at the parameters specific to only to the subset which means suppose I am doing the you know the whole blood right, I am doing a whole blood analysis right and I have collected this as like FL1 versus or suppose site scatter versus FSC right. So I am going to get this right and I am not interested in all these, I am only interested in macrophages. So if I go with this particular type of gate, it is only going to give you the information about the macrophages or lymphocytes or lymphocytes. So all these different types of cell type can be collected and then you only get the information about that. Although you are doing the complete proved sample, you are doing the blood analysis, but it only going to give you the information about the macrophage or the lymphocytes.

And you can actually have the different types of gates, you can have the rectangular

gate, you can have elliptical gate, you can have polygons, quadrants, histogram and so on. So to explain you these things in more detail with what you know with an instrument, we have prepared a small demo clip where the students are actually going to explain you all these steps in more detail. Hello everyone, in this video we will be discussing about flow cytometry, its equipment and the software related to it. Flow cytometry is a basic technique in which the chemical or physical characteristics of a cell or a population of cells are determined by the instrument.

In this process, the cells are suspended in a fluid, mostly saline, 0.9% saline and is passed through a beam of light and then the physical and chemical properties are recorded. Coming to the equipment, this is a standard flow cytometer equipment manufactured by BD Biosensors, BD Fax Caliber. So basically there are 3 parts in this machine, the fluidic chamber, the sample induction port and the fluidic panel. In the fluidic drawer, there are basically 2 tanks.

This tank is used known as the sheet fluid tank and this is a wastage reservoir. In this one, we have to pour 0.9% saline which is used as the sheet fluid and this is which passes to the sample and then on to the detector. Coming to the sample induction port, in sample induction port only certain types of tubes are known, are used which are known as the Fax tubes. In this sample injection port, we have to change the sample like this. Coming to the fluidic panel, as you can see there are 6 buttons, the low, medium, high, run, standby and prime.

We have to always remember that whenever we are changing the sample, the machine should always be on standby and when we are acquiring the data, it should be on run. The low, medium and high buttons represent the speed or the speed with which the machine sucks the sample. For low, the machine sucks the sample at 12 microliters per minute, for medium at around 35 microliters per minute and for high around 60 microliters per minute. This prime button is used when for example, if the sample injection port is stuck with air bubbles or is stuck with like samples like which are having doublets or debris. So, in order to analyze the data of the Fax, we use the software known as the CellQuest Pro.

We can find its icon in the toolbar as shown in the window. After we open the CellQuest Pro software, the first thing we need to do is acquire and then connect it to cytometer. After connecting it to cytometer, two windows pop up, the acquisition control and the browser and title. In the browser and title, we can save the data as well as the data file. This acquisition control is used to acquire the data and set the setup button. When the setup button is on low, we can try and error the data in this main panel and when we remove the setup and acquire the data, only then the data will be saved.

One more thing we have to remember is after connecting to the cytometer, we should make sure that in the instrument, the tank is pressurized and it should not be depressurized. So, when the tank is pressurized, only then the sheath fluid will flow through the instrument. Then coming to the things required for acquiring the data, we need a parameter description, we need counters, we need detector and amps, we need status. So, coming to the detector and amps, as we can see that there are multiple detectors like FSC, SSC, FL1, FL2, FL3 and FL4.

The FSC and SSC are related to the forward scattering and the side scattering. The forward scattering tells us about the cell size, like how large the cell is and the side scattering tells us the complexity or the granularity of the cell. The FL1 channel is for the green fluorescence, FL2 is for red fluorescence, FL3 is for red or orange and FL4 is for blue fluorescence. As we can see in the voltage, we can see that there are 5 voltage gates like E00, E01, E02, E03 and E-1. For cancer cells, we mostly use E-1 because the cell size is large and as the cell size decreases, the E01, E02 increases. For example, the E01, E02, E03, these are used for smaller cells such as RBCs or microbes or macrophages.

We can control the voltage using this tool bar. I will show you how to do that when we acquire the data. In the status panel, we can see that the status is showing standby, which means that we are not yet acquiring the data and the system is on standby. In the counters, we can see that the total events per second is 0 and the events per second is 0 because when we start, it will change. Coming to the plots, the most basic plot in FACTS is the dot plot.

No data is valid without the dot plot. As soon as we plot the analysis dot plot, the inspector dot plot window pops up. In this, we have to click at the corner of the analysis dot plot and then in the plot type, we have to do acquisition and analysis. And then we can change the X parameter or the Y parameter using one of these options. For this demo, we are just showing how to acquire the data using a cancer cell line.

Now, I will show how to acquire the data. Before acquiring the data, we should make sure that how many number of events we want to record before stopping. The default is set to 10,000. So we will reduce it to 5,000 just in case just to see the data how it is going and then click okay. And after that, on the fluidics panel, we have to press run in order for the machine to set the sample from the FACTS tube.

As soon as we press the run button, we should press acquire and then only we can see the data. As you can see that, we can see that there are multiple cells being shown here

in the near the 00 and they are showing away from the access as well. So we can change the direction of the flow using this according to our requirement in order to set the population. So let's say that I want to stop the data after 5,000 events. So before that, we need to save the data. So let's say that we are going to save the data in this file and then the data file can be written as, then click okay.

And if we remove from the setup, then we can actually save the data. As we can see here, the total events are being recorded like till now 500, 600 events are recorded and the events per second are 75. As we have seen on the fluidics panel that there are three buttons, the low, medium and high. This actually decide how many events are being recorded per second. For example, if we press the low, then the events per second will be recorded low because we are actually taking low amount of sample from the fax tube.

And if you press high, the number will increase. It also depends on the concentration of the sample. So mostly if the concentration is high, we should opt the low button in the fluidics panel so that the complete number of events will, so that the instrument can record complete number of events in a right manner. Now as we can see that as we have sent the total events to 5,000, after 5,000, the data looks like this. So the machine stops recording the data after 5,000 events because we have set to that. So after that, if you want, like this one, let's say that if this was the untreated and we have acquired the data of the untreated at a particular voltage and amp gain in order to acquire the data of the treated samples, we should not change these parameters or else we will not know what is actually difference between the untreated and the treated samples.

So after this, let's say that we have completed the experiment. So after this, we need to watch the system so that the next person or when we use the next time, it will be easy for us to operate. So for that, we have to clean the system with 2% sodium hypochlorite solution and we have to remove from the setup so that it does not save the data. Now I have changed the sample from the sample to the 2% sodium hypochlorite. As we can see that some of the cells will still be there in the sample induction port and that needs to be cleared.

So for that purpose, we are washing the system with 2% sodium hypochlorite at high pressure. We have to keep washing the system till the events per second remain 14 or 15 for a longer time. So while handling the FACTS instrument, there are some precautions to be taken. For example, when we change the sample in the sample injection port, we should always make sure that the system is on the standby mode and also when we are done analyzing the data, the system should always be in the standby mode and after using the instrument, we should always depressurize the sheet fluid tank because if it is not depressurized, it may actually harm the system. Now we are going to see how to

process the raw data which we have just acquired in the FCS Express Data software.

So far we have seen how to acquire the data in the FACTS equipment. After we acquire the data in the FACTS equipment, we have to process the data in the software known as the FCS Express 5 flow software. So to begin with, after we open the software, we open the new layout. After we open the new layout, a window appears in which variable options are there such as home, insert, gating, batch, format, text, data, multi-cycle view. So in the home tab, we can see that we can take a new page according to our needs like blank title according to our requirements and also we can check the layout. We can also take the layout which is best suitable for our plots and we can also change the size of the plot with respect to our requirement and if you want to change the orientation, we can use

portrait	or	landscape
----------	----	-----------

And also in the insert, we can use whichever data, whichever plot which we want to use, for example histogram or multi-cycle DNA or proliferation or dot, density, color, contour, scatter, etc. So the most basic plot is the dot plot which is essential to every facts data. So after that, we can also do gating. So gating is a mechanism in which we can segregate two or more different populations from a given population of cells.

For example, if you want to segregate the debris from the singlet and the doublet cells, we can use gating. So gating can be of different types like ellipse, rectangle, polygon free form. And also we can do the quadrant analysis. For example, in acridine orange, propidium iodide or annexin 5 propidium iodide staining for the live and dead cells, we can use quadrant analysis in which the first quadrant shows the live cells and so on. In the batch, in the format, we can select which format which we want to use.

Anyhow, all the options will be enabled once we load the data. So let's load the data and see how it looks. So after we input the file into the software, a window appears in which we can use different types of plots such as dot, density, color, contour, etc. So the dot, the density, color dot, contour, these all and the scatter plot, these all are the same plots but in a different way of representation. So let's see how does that look. So for example, I have taken an untreated sample obtained on a FL2A channel.

So FL2A channel is responsible for red fluorescence. So we can see that this is the multi-cycle DNA plot and this is the histogram plot as we have taken it on the FL2 channel. So we will switch to FL2A and this is the scatter plot. This is the density, the contour plot. This is a color plot and this is a density plot and the final one is the color plot.

So the first one is the FSC versus SSC. This is a dot plot. This is a basic dot plot which

represents how the data looks. For example, as the population of the cells goes away from the 00 mark, we can say that the granularity as well as the size of the cells increase. So if you want to exclude that from the data, we can exclude that using gating option. Also as we can see that the second plot which is the density plot, it is almost as same as the dot plot but with respect to the density. For example, the blue color in the middle represents a denser population and as it goes outward, the population of the cells decrease.

So the same dot plot can be represented in a contour plot as shown in the third plot and the fourth plot is the histogram in which we can see the difference, the variable intensities for the different cell cycle phases and the fifth plot is the DNA cell cycle plots in which we can see different phases of cell cycle like G1 as G2. The scatter plot is just another representation of the dot plot in which the events are just highlighted in a respectable way. And also the last plot is a color plot. Actually in a color plot, we can see two different populations assigned with a different color but because we have used only one dye, so only one color is visible. So in this way, we can check the different plots and also if you want to see the statistics and also the details of the plot, we just have to click right and then press the format option and we can compensate the data, we can cut, copy, paste.

And in the format option, we can actually do a variable number of things like changing the border color, overlaying the data, the dot options, the size, the background, the axis. We can do a whole lot of options which will be helpful for us in order to present the data in a more appropriate manner. So this is all with respect to the FCS-CFI Express software.

Hopefully, this was helpful for everyone. So with this, I would like to conclude my lecture here. In our subsequent lecture, we are going to study or we are going to discuss some more aspects related to the biological system. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 02
Basics of Biological system (Part 2)
Lecture-09 Cell Cycle and Control (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati. And what we were discussing, we were discussing about the basics of the biological system and in this context so far what we have discussed, we have discussed about the cellular structures. So, initially we discuss about the prokaryotic structures followed by the eukaryotic structures and when we were discussing about the eukaryotic structures, we have also discussed about the different organelles what are present in the eukaryotic cells and how you can be able to separate them with the help of different types of fractionation techniques. We have discussed about the density gradient identification and as well as the differential certifications and in the previous lectures, we have also discussed about the cellular metabolism. So, we have discussed about the anabolic reactions and we also discuss about the catabolic reactions.

So, within the catabolic reaction, we discuss about the carbohydrate metabolisms and lipid metabolisms whereas within the anabolic reactions, we discuss about the synthesis of the different types of amino acids and so on. So, in today's lecture, what we are going to do is we are going to discuss about the how the cell is utilizing this energy for different types of activities and one of such different activity which is very crucial for the cell is the cell division and the cell cycle. So, I hope you have understood the how the flow cytometry is actually analyzing the different types of cells and then you can actually be able to exploit this particular technique not only for studying the cell cycle, but also for the other kinds of applications. So, these are the some of the classical flows and dye what you can actually use into the flow cytometry.

You can actually use the FITC, PE, per CP, APC, Texas red, you can actually use the PE Texas red and PE silane and these are the emission wavelength. So, all most of these emission wavelengths are available in the and they will actually go into the very clear cut you know separation from each other so that you can be able to use them either singly or in a combinations. So, now come back to the our original question how you can be able to use the flow cytometry to study the cell cycle analysis. So, for the cell cycle analysis what we have is we require to stain the cells for the DNA binding dye. So, one of such DNA binding dye is the propodium iodide.

So, if you stain the cells with propodium iodide, so this is the propodium iodide structure and it is actually going to stain the DNA by the you know by using this positively charged residues. So, the propidium iodide is a flow-set dye frequently employed in the analysis of cell cycle to label the DNA. PI has a unique property of being unable to enter the live cell but capable of permeating the cell membrane of the seeds or the diet cells through its ability to DNA and subsequent examination of red flow set upon excitation with a specific wavelength. PI enables the evaluation of individual cells DNA content and facilitate the examination of their distribution within the cell cycle. Remember that what we are supposed to study we are supposed to study how which cell has the 2x amount of DNA and which cell has the 1x amount of DNA.

Apart from that it is also going to combine this information with the size and they using these two property you can be able to study the whole cell cycle. Remember that G1 phase the amount of DNA is going to be 1x and size is also going to be small whereas S phase you are going to have the 2x whereas G2 phase you are going to have this bigger size and as well as the 2x amount of DNA. So, that is a way in which you can be able to use the combination and you can be able to study the cell cycle using that. So, if you want to do the cell cycle analysis these are the material what you require. So, I am not going to go through with the content right.

So, you require the well plate, propionate iodide, RNAs and all that and this is the protocol what you are going to do. So, you take the cell in a 6 well plate you on the day of experiment wash these cells and then treat the cell with desired concentration and incubate. So, mostly people do the cell cycle analysis when they are doing the treatment right. For example, the cancer cell was if they are under the treatment their cell cycle maybe get disabulated and that is how they will actually going to undergo the apoptosis or cell death. So, for those kind of experiment what you are going to do is you are going to treat the cells with a particular anti-cancer agents or compound and then you are going to study whether it is affecting the cell cycle or not.

So, for example, after the 24 of treatment remove the media and scrap the cell with a 16 or 0.6 percent ETTA, centrifuge at low rpm and discard supernatant using the 16 and retain the cells as a pellet. Then you add the 250 microlitre of EBS to resuspend the pellet cell and keep the suspension on ice for a couple of minute, add 1 ml of 70 percent ethanol and under work test condition for the fixation of the cell transfer the sample from ice to minus 20 for 3 to 4 hours. And then you are going to have the incubation with the 70 percent ethanol centrifuge at 5000 rpm discard supernatant and save the pellet then you add 1 ml of ice cold PBS and gently resuspend the cell centrifuge the sample at 5000 rpm at 4 degree for 10 minutes to remove any residual media and debris remaining from the cell. So, you take the washing step with the PBS two more times after washing the

cells with PBS add the RNAs so that you are actually going to remove all the samples all the RNA set signal actually because RNA signal is going to be very high because it is present in the cytosol.

So, you do not want that kind of disturbance. So, with RNAs one where the final condition would be 20 to 30 microgram and incubate at 37 degree Celsius for in water bath for 2 hours. After the RNA degradation by the RNAs a stain the cells with propodium iodide final consideration is 50 to 70 microgram for 30 minutes before acquiring the data flow cytometry. And when you are doing the data acquisitions, analyze the stain cell using the flow cytometer equipped with the appropriate filters for the propidium iodide. Usually there is an excitation at 488 and the emission at the 585 nanometer.

So, you should use slice the standard flow cytometer where you can have the different these type of lasers in a you know the detectors. Adjust the flow cytometer setting for appropriate sources and forward and side scattered. Run the stain cells on a flow cytometer collect the data for at least 10,000 events. So, every event means the cell actually. So, you have to collect 10,000 cells data.

Analyze the acquired data using the flow cytometry software, plot the DNA content and onto the x-axis and the cell count or the event onto the y-axis. The resulting histogram will typically show the distinct peak in the different phases of the cell cycle such as G1, S and the G2. Data analysis, the data analysis can be done with the FCS-5 XPREN software. These are the some of the precautions what you are supposed to take and to avoid any disruption in the distribution of cell population is recommended not to seed or treat the cells when they reach the confluence between 90 and 100. So, you should not take a very old cell.

So, it should be cells which are under the actively dividing cells should not be like confluentish because that the cells are going to be old and they will not be dividing. To avoid obtaining unwanted background signal, it is important to run the cells labeled with the propidium iodide in a PBS when using a flow cytometry instead of using the complete or the incomplete media. Ensure to avoid excessive pipetting when the scraping the cells from the plant because that is actually going to damage the cells. So to explain you how to perform these and how you can be able to study the cell cycle using the flow cytometry, we have prepared a small demo where I will take you to my lab and where the students are actually going to explain you how you can be able to stain the cells, how you are going to acquire the sample and so on. Hello everyone, in this video we will be discussing about how to perform cell cycle analysis on a FACS equipment.

So basically the cell cycle analysis gives us the details about how much population of a given number of cells is present in which phase of the cell cycle. For example, the G0 or the G1 phase, the synthesis phase or the G2 or M phase. So we can design the experiment according to our needs. For example, if we want to find out how a chemical compound inhibits a certain population of cells, we can do a time dependent experiment or also we can do a concentration dependent experiment. So in both the ways we can find out which particular phase of the cycle is being inhibited.

The procedure for cell cycle analysis is very simple. First we trypsinize the cells from 100 mm cell culture dish and seed approximately 1 million cells in a 6-well plate. After the cells are seeded, we incubate it for 12 to 14 hours so that the cells get adhered completely to the dish. After the cells are adhered, we give the appropriate treatment but after washing it twice with PBS. After the appropriate amount of treatment, let's say 12 hours or 24 hours or 48 hours, we trypsinize the cells and collect the pellet.

We wash the pellet two times with PBS and then centrifuge again and then resuspend the pellet in 1 ml of 70% ethanol. And then this mixture is kept in minus 20 degrees for 12 hours for fixating the cells. After the cells are fixed with 70% ethanol, we wash the cells with 5 ml of PBS and then collect the pellet and resuspend the pellet in 1 ml of PBS. Then we provide the appropriate RNA treatment to remove all the RNA from the mixture because RNA, if present in the mixture, might interfere with our cell cycle analysis. After incubating the sample with RNAs for 2 hours, we give the appropriate Propeidium iodide treatment.

The working concentration for Propeidium iodide in cell cycle analysis is 50 microgram per ml. After giving the treatment with Propeidium iodide for around 1 hour, we analyze the cells in fax equipment. After the samples are prepared, we need to analyze the data in the fax equipment. The first thing we do is open the cell Quest Pro software. After that, we connect it to the cytometer.

After connecting to the cytometer, we need shooting such as counters, detector and amps, status, and also we need the dot blot, acquisition and analysis. We need another dot blot for FL-2A and SSC because in the FL-2A channel, the Propeidium iodide emits red fluorescence. We also need a histogram to see the cell cycle phases. For that purpose, we need the FL-2A channel. We need another dot blot in order to see whether there is a presence of any doublets.

There is a chance that there might be a presence of the doublets in the data like for example the clumps of cells, so we have to exclude that from the data. For that purpose, we take the FL-2A and FL-2W channel. Now that everything is ready, now we can

change the directory for acquiring the data. We can also choose the file name. So we will name it as unwitted1 and we will keep the file count to 1.

For checking the data, we have kept it on the setup so that we can see whether the data is coming properly or not and then we can remove it from the setup and then acquire the data. So now we are going to load the sample onto the sample injection port and then press acquire. As we can see that the number of events have started recording. In the FSC and the SSC plot, we can see that most of the population is seen between the 200 and the 200 mark. The population which is coming away from the doublets.

In FL-2A and SSC, we can see that there are three different populations, G1, S and G2. And in the FL-2A and the counts, we can see that the longest peak is the G1, the between one is the S and the smallest peak is the G2. So before acquiring the data, we have to set the number of events. So we go to acquisition and storage and press record 10,000 events. So in order to record the data, we have to remove it from the setup and then setup and then press acquire.

In the FL-2 and the FL-2W channel, we can see that we can see there are two different populations. One in line and another is the presence of some doublet or clumps of cells. This dating population might be presence of the doublet cells or clumps of cells. So while acquiring the data, we can see this but in the FCS Express software, we can remove that. So now that we have acquired the data, now we can take the data from the untreated sample.

We have to remember that we don't have to change the parameter conditions in order to compare the untreated and the treated samples. Now we have to change the sample name from untreated to treated and also change the file count to 1. Press ok and then acquire. We can see that there is a little shift in the S and the G2 phase but we can't say for sure that there is a change in the data. So for that, we have to acquire, we have to process the data on the FCS Express software in order to see what is the difference between the untreated and the treated sample.

So now we will remove it from the setup and then acquire the data for the treated sample. Sometimes we can see that the events per second will be low. So in order to increase the events per second, we can pause the recording and then tap the sample once or twice in a while to shape the sample and then the flow will be continuous again from the sample injection port and the events per second will increase. Now we can see that the events per second has increased because we have tapped the sample and the data acquisition will be little bit faster. So now we have acquired the data for the untreated and the treated samples.

In order to see in which phase of the cell cycle the arrest has taken place, we have to process the data in the FCS Express 5 software. So now that we have acquired the data on the FACS equipment, we have to now process the data in order to see the difference between the untreated and the treated sample. For that purpose, we use the FCS Express 5 flow software in order to process the data. So for that purpose, we use the new layout and then change the mode to landscape because it is easier to work in the landscape. And then after that, we go to the data in the toolbar and press open and then go to the folder where our data is saved and press open the untreated file.

As we can see that there are multiple options available, dot density, color, dot contour, surface histogram and multi-cycle DNA and kinetics. For this cell cycle DNA, we only need multi-cycle DNA plot and the dot plot. So we are going to open these two. Because we have recorded our data on the FL2A channel and also because the propidium iodide is only shown in the FL2A channel, we are going to open the FL2A channel. So now after we have opened the dot plot and the cell cycle DNA plot, we can see the, means how much population of cells is represented by the G1s and G2 by right clicking on the plot and then selecting statistics and then show DNA cycle statistics.

So a small window will appear in which it will show that the %G1, %G2 and %s. So this is for the untreated sample. Let's see for the treated sample. As we can see that there is some change in the untreated and the treated sample both in the both plots, the dot plot and as well as the cell cycle plot. So in order to see how much % of the cell cycle phases have changed, we right click on the plot, then show DNA cycle statistics and then place it right beside the untreated one in order to compare.

So we can see that there has been a reduction in the G1 phase from 61 to 49 and from G2 also from 15 to 12. But there has been an increase in the %s phase which has increased from 23 to 38. So we can say that there is a significant change in the phases of the cell cycle. But we can only be sure after doing the experiment in the triplicate so that we get the proper standard deviation and also the standard error.

So this is one way of processing the data. Another way of processing data is by gating. As we can see in the FSC and the SSC channel that there are some debris near the 00 point and also there are some population of cells which are very far from the 00 representing clumps of cell or may be doublets which we have to take care using gating technique. So in order to explain the gating, I will use another page and then open the data and then press OK. So for gating we need multiple plots. The first one is FSC and SSC and the second one is between FL2A and SSC and the third one is between FL2A and FL2W.

The fourth one is between FL2A and FL2H and the last one is between FL2H and FL2W. So we have to start the gating from the first plot, the FSC and the SSC. So in order to do the gating, we have to go to the gating option in the toolbar and then we can choose any one like ellipse or the rectangle, polygon, freeform. These are the shapes of the gating. So we will go with the polygon gating because it is easier to handle and then we can get our required population using polygon gating tool and then just select the required population.

So here we are excluding the debris and the clumps of cells which might be interfering with our actual data and then we can select what gate colour we can give and also rename the gate 1 to our particular type but we will go with the standard one and then press OK. And then we have to apply this gating to the second plot. We can just drag and drop it on the second plot. So we can see here that we have excluded some population. So but still there is some population which can still interfere with our data.

So we will do the gating again and then we will exclude some more population which might not be helpful to us. And then we will name this as gate 2. So in order to apply this one, this gate to the third plot, just drag it and drop and then we can see that another number of population have reduced and in the third plot as well we can reduce this population which is a little bit distinct from the singlet cells. So we have taken another gating using polygon and then we are only now going to select the singlet cells and then we will name this as gate 3. From gate 3 onwards we just have to apply the gates to the next plot and then we can apply the plot, we can apply the gate to the plots using just by just selecting the plot and then going to the top left corner and then selecting the gate.

So as we have seen that we have reduced significant number of population and we have only selected the population which might be helpful to process the data. And then finally we apply the gate 3 which is the final gate and then we will see the change in the cell cycle statistic plot. So this is for the untreated one. In order to show the cell cycle statistics just press statistics and then show DNA cycle statistics. As we can see in the ungated one we have seen that the percentage of G1 was 61 whereas in the gated one it is 67.

1 and the percentage G2 is 8.9 and the percentage S is 23. Similarly we can do for the rated one but we don't have to follow the whole procedure. We can just copy all the plots to a new page and then press paste and then go to data. And then select the treated one. It will just replace all the plots with the treated data. So but the treated data and the untreated data is little bit different so we just have to move the gate.

So in order to see the cell cycle statistics for the treated one we can select the statistics DNA cycle statistics and then we get the DNA cycle statistics. So in this way we can get the process the data for the untreated and the treated sample in the cell cycle DNA statistics using gating as well as non-gating technique. So hopefully this video was helpful for everyone. So at the end of this experiment what we are going to do is what you are going to see is you are going to see all the cell cycle stages.

You are going to see the G1. You are going to see the G2. You are going to see M. You are going to see the S phases. And all these are going to be analyzed by a flow cytometry analysis software. You can use any standard software what is available into your laboratory.

You are not bound to use only the FCS-5 receptor. And what will happen is that when you analyze that it is actually going to give you the different phases. So when you plot the FL2 versus the number of cells this is the unfiltered sample and this is the sample which has been treated with the anti-cancer compound. So what you see here is that this is the G1 phase right. This is the G2 phase and this is the other phases right.

So what you see here is that this is going to be the G1 phase. This is going to be the G2 by M phase and this is going to be the S phase right. So what you see here is this. This is the S phase okay. These are the cells what are present in the S phase.

And these are the cells that are present in the G1 phase. And these are the cells what are present you see this red color right. These are the cells which are present in the G2 by M phase. So although the separation of the M phase from the G2 is very very different because they both have the same amount of DNA and they also have the same size. Because at the end of the M phase you are going to have the 1X of DNA right.

But before that it is actually going to have the 2X of DNA. So that is why it is very difficult to separate out the G2 by G2 phases G2 phase from the M phase but it will actually going to give you an idea right. So this is the what it says is that out of the total number of cells the 54% cells are under the G1 phase, 6.72% cells are into the G2 phase and only the 39% cells are under the G2 by G1 phase. What you see here also that S phase is actually the phase which is responsible for the DNA synthesis right.

And what you see here is that the S phase is now 20%. This means there is a disturbance within the S phase and that is how it is actually going to suppress the growth of the cell right. And these are the some of the information what you are going to get when you are going to do the analysis of the facts analysis. Apart from the flow cytometry you can also be able to do the traditional methods where you can actually be

able to you know prepare you can take the plant tissue and you can actually be able to extract the chromosomes and you can actually be able to visualize the cells and you can be able to study the different types of phases using the other methods. And in this the advantage is that you can not only be able to study the G₁, G₂ and S phase you can also be able to study different phases within the mitosis such as interphase, you know interphase, metaphase, telophase and anaphase.

So this is what we are going to discuss now. So if you want to study the M phase the flow cytometry is not good for studying the M phases right or different phases within the M phase okay. So for that we are actually going to prepare the chromosome samples and it is actually going to give you the different sample. So mitosis and meiosis you are actually going to prepare the chromosome preparations okay and that chromosome preparation can give you the clear idea about what are the different sets of cells present in the mitosis and within the mitosis how many cells are present in the interface, how many cells are present in the metaphase, anaphase and telophase. And for doing this what you require is you require a plant sample so you can actually be able to do this in a root tip right and you require a plant cell samples, you require the if you want to study the same thing in the meiosis then you require the flower beds, a flower bud right because flower bud is actually going to be actively divided by the mitotic phases and then you require all these material. Apart from that you require a compound microscope and you also require a testiminator for storing the prepared reagents.

You also have to prepare the different types of dyes such as the acetone or the cetin dye, you also require the coronary solutions and so on. You require so how you are going to prepare all these solutions this is all the recipe is already been given and then you require the two different types of methods. So you can actually be able to prepare the mitotic chromosome preparations. So mitotic chromosome preparation this is the lengthy procedure and we are actually going to you know so these are the methods what you are actually going to follow. So this is the protocol if you follow this step by step it is actually going to give you the chromosome preparations okay.

And that and then apart from that if you want to study the meiosis you can actually be able to prepare the chromosome preparation in a different way. You are actually going to use the flower buds if you want to study the meiosis whereas if you want to study the mitosis you are actually going to use the root tips or the somatic tissues. And at the end of this what you are going to do is you are going to prepare you are going to take some precautions such as you are going to use the, collect the plant material which is you know which should be performed in a bright sunny day to ensure the proper mitotic stage of the development. All the collection timing may vary depend in between the plants. Cloudy rainy should be avoided for the material collection and use of the freshly prepared

solution of the acetooxygen solution should be used during the preparation of the slides.

So I have prepared a small demo clip into one of my colleagues lab and there the students are actually going to show you how you can be able to prepare the chromosomal preps for studying the mitosis and meiosis. Hello everyone I am Rajendra PhD student from IIT Guwahati. In this tutorial I am going to demonstrate you the practical aspects of mitosis cell division. In this method we will explain you each and every details of mitosis starting from sample preparations to the microscopic observation.

We will explain you each and every details. In today's demonstration somatic chromosomes will be studied from onion root tips for mitosis and flower buds will be used for meiosis studies. Various chemicals and materials will be required such as 0.02% 8-hydroxyquinoline, carnois solution, Glacial acetic acid 45%, 1 normal hydrochloric acid, 2% acetoarsin, wash glass, cover slip and cover glass, burner, glass beaker, pipettes, various size of the forceps, blades, filter papers or the blotting sheets and ependop. Root tips from healthy onion plants were collected at 9 am in morning.

Initially epical region of the root tips were cut by using surf blade. 2-3 pieces of root tips 1 cm in length were selected. 8-hydroxyquinoline 0.02% were previously prepared and kept in amber glass bottle. 1 ml of 8-HQ previously kept in ependop tubes were used and those cutted root tips were kept inside ependop tube. After that this ependop tube will be stored in refrigerator at 4 degree Celsius temperature for 4 hours.

This whole process is called pretreatment. After 4 hours of pretreatment at 4 degree Celsius temperature the 8-hydroxyquinoline has been discarded. Therefore, the pretreated root tips now will be fixed in carnois solution. Carnois solution containing absolute alcohol, chloroform and glacial acetic acid at 6 x 3 x 1 ratio. The ependop tube containing 1 ml of carnois solutions with pretreated root tips now will be stored at room temperature for 48 hours.

After fixation in carnois solution for 48 hours that solution has been removed. Therefore, that fixed root tips has been transferred in vase glass followed by addition of one normal hydrochloric acid. one drop and 2% aceto oresin. The mixture of aceto oresin and hydrochloric acid has been heated gently over the burner. This process is called staining.

Now the stained root tips has been placed over the glass slide. Lower portion of the root tips has been removed by sharp blade. Only the meristematic region has been taken for analysis. Root tips were then squashed and mounted by cover slip. This process is very careful to prevent entering of air bubbles. Now the sample was covered by the filter

paper and gradual pressure has been applied by fingers in order to spread out the cells.

Now the slide is ready for observation under microscope. Now the slide has been placed under microscope. Photomicrographs were taken with Carl Zayes microscope having 10x, 20x, 40x, 60x objective lenses. We can see here from the microscopic field that chromosomes of onion root tip cells is clearly visible along with cell divisional stages like anaphases and telophase also visible. In this method we have explained each and every steps of mitosis starting from sample preparation to microscopic observation. Hope this video will help you to prepare the slide of any plant sample for the study of mitosis.

The meiotic cell division process starting from sample preparation to microscopic observation. We will explain you each and every steps in detail. Flower buds of onions will be used for mitosis study. Initially flower buds were collected during the flowering season in morning between 11 to 11.30 am and has been fixed in carnoise solution containing absolute alcohol, chloroform and glacial acetic acid 6H to 3H to 1 ratio.

Those fixed sample has been placed over a watch glass and a single flower bud has been selected for smear process. Flower buds size 1mm in length has been taken and placed over a surface of glass slide. After that sepal and petals were removed initially from that selected flower buds. Therefore the anthers were removed.

Special isolated anthers are clearly visible on the surface of glass slide. One drop of 2% acetoarsin has been given over the anthers. With the help of iron needles the anthers were ruptured and pollen mother cell has been released. Very gentle pressures has been applied over the anthers. This process is called smear technique. Now the anthers walls has been removed so that we can observe the various stages of pollen mother cells that is undergoing meiosis.

After that one cover slip has been placed over the sample with the help of pointed iron needle. This process needs extra precautions to prevent the entry of any air bubbles in between slide and cover slips. We can use blotting sheets or filter paper to remove excess stains. Now the slide is ready for microscopic observation. After that the slide has been placed under microscopes to capture photomicrographs. Different objective lenses has been used such as 10x, 20x, 40x and 60x for capturing the various stages of meiotic cell division.

Here we can see some earlier stages of pollen mother cells. In this image we can see after second meiotic division two cell stages is formed but the four distinct nuclei has been reached in two different poles. In this process we have explained each and every

step of meiosis cell division process starting from sample preparations to the microscopic observations. Hope these videos will help you to prepare the slides from any flower buds in your plant sample. Thank you for listening. So, this is all about the cell cycle and let us discuss about what is the role of cell cycle into the different properties or different section of the biological function of the cell.

So, the role of cell cycle is actually required for development of the growth. So, the development of cell cycle single cell into the multicellular system is possible due to the cell cycle and division. Then it also requires for the cell replacement. So, eukaryotic cells have the predefined lifespan after that period it needs to be replaced with the new one. It is possible due to cell division and making more cellular properties.

For example, the human RBC has a lifespan of three months. New RBCs are formed from the bone marrow by the cell division. Then it also requires for the regeneration. So, cellular damages and injury is an integral part of the living system and the cell division is the primary event required for the synthesis of the lost or the damaged organ. Then it is a very very important method for the asexual reproduction. So asexual reproduction is common in the lower invertebrates such as and in these organisms the cell divide to form the new cells and these newly formed cells give rise to the new organism for example, hydra or planaria or amoeba.

Now the last part is that how you are actually going to control the cell cycle because if it is a cell division it actually increases the cell mass it has to be very precisely controlled. So, cell cycle is controlled at the multiple stages. It is actually going to be controlled at the interface of the G1 to S phase. It is going to be controlled as the S to G2 phase and also going to be controlled at the G2 to M phase and all that. So at this stage when the cell is entering into the G1 phase it is going to be checked for different types of parameters so that it should not happen spontaneously.

So cell cycle at different step is tightly controlled by the cell cycle check proteins. The cell cycle check proteins are used to ensure the completion of different steps and repair of the DNA damage. The main checkpoints are present at the G1 to S phase, G2 to M phase and the M phase. Each checkpoint is controlled by the mutual interaction between the cyclin protein and the cyclin dependent protein kinase or CDKs. P53 protein are known to control many events through the G1, S and G2M checkpoints.

So what will happen if there will be a dysregulation or there will be a control gone right. So what will happen if these events goes wrong right. Dysregulation of the cell cycle and control mechanism give rise to the enormous growth of that particular cell and that is nothing but the tumor actually. Under certain number of cell division every cell will enter

into the G0 phase and ceases its cell division. In that case of tumors cells lost the control mechanism and multiply indefinitely to give rise to the cell mass.

These cells are taking nutrition but not performing the functions. Retinoblastoma or the RB proteins, P53 are the crucial factors cellular factor responsible for the cell cycle control and play crucial role into the tumor deployment. If you want to study more or if you want to study more about the cell cycle control, I have given you this reference and you can be able to study or you can actually be able to go through with the content. So what we have discussed, we have discussed about the cell cycle and control. We have discussed about the different phases of the cell cycle whether it is G1 phase, G2 phase, M and S phase and we also discuss about the relevance of these phases.

Apart from that we have discussed about the two different methods. One is the traditional method where you are actually going to prepare the chromosomal preps to study the mitotic phases and the other is the more advanced technique of the flow cytometry where you can actually be able to separate the cells based on the DNA content at the size and both of these techniques are going to be very robust to study the cell cycle and the different phases of the cell cycle. So with this I would like to conclude my lecture here in our subsequent lecture we are going to study or we are going to discuss some more aspects related to the biological system. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 02
Basics of Biological system (Part 2)
Lecture-10 Program Cell Death

Hello everyone! This is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati and what we were discussing we were discussing about the different aspects of the cells or the biological system. In this context so far what we have discussed we have discussed about the structure of the cells whether it is a prokaryotic cell or the eukaryotic cell. So in the current module what we are discussing we are discussing about the cellular metabolism so we have discussed about the carbohydrate metabolisms, lipid metabolisms and in addition when we have also discussed about the protein synthesis and so all the purpose of these catabolic reaction is to generate the energy and then this energy is actually going to be utilized for the synthesis of the new biomolecules and if you see that the synthesis of these new biomolecules is directly linked to the growth of that particular organism and subsequent to that it is actually going to result into the division of these cells and so that the number of cells are actually going to increase and this is what we have discussed in the previous lecture. In the previous lecture we have discussed about the cell division and how the different phases are present in the cell division during the and we have discussed about the interface we have discussed about the S phase G1 and G2 and as well as the M phases and in the previous lecture we have also seen how you can be able to how you can be able to experimentally verify these cells in the different stages whether it is mitosis, meiosis or the different cell cycles studied by the effects flow cytometry. In the current lecture what we are going to discuss we are going to discuss about the programmed cell death.

So as the name suggests this is a this these are the set of set of the cascades of reactions which are going to be governed by the molecular players and they are responsible for the death of the cell in a very very systematic way so that it should not cause any harm to the organisms. Now first let's discuss about what is the life cycle of a cell. So if you think about the life cycle of a cell it starts with a cell okay whether it is a prokaryotic cell or the eukaryotic cell and when the cell take up the nutrition from the outside it is actually going to take this nutrition and generate the energy right this energy it is going to be generated by the catabolic reactions and we have there we have discussed about the carbohydrate metabolism and as well as the lipid metabolisms this energy is actually going to be used for many purposes such as this is going to be used for the growth of the cell and it is also going to be used for the reproduction right because every cell wants to increase its number so it's going to be used for the reproduction of the cell. Now this will

continue until the uptake of the nutrition and the production of energy would be on a higher side compared to the energy what it requires for the other kinds of processes and after some time when the cell will go through a process of aging what happen is that this cell will actually go to the different types of changes right.

One of the things one of the serious change is that it's actually going to enter into a non dividing phase which is called as G0 phase and once it enters into the G0 phase it will stop the division right so it will not going to produce new cells and on the other hand it is actually going to just maintain the basal level of activities and as a result it is not going to perform many functions. After this it will enter into another phase which is called as death phase right because the every every cell has its definite lifespan so it is actually going to enter into the death and the death within the cell can be induced by the two different processes. It can be done either by a programmed method or a programmed manner which is called as programmed cell death or apoptosis or it can be done in a another method which is called as necrosis okay. We are going to discuss in detail about the differences between the Bogram cell death apoptosis as well as the necrosis and what are the contrasting feature of these two things so and either of these methods are responsible for the death of the cell right and the cell has to take a decision whether it wants to go for the apoptotic pathway or the necrotic pathway. So in today's lecture what we are going to discuss we are going to discuss what actually induces the cell to go for a suicide pathway or apoptotic pathway.

What is the definition of the apoptosis and its major features right so what are the different hallmarks of the apoptosis what are the different events are happening and then the steps involved in the apoptosis difference between the apoptosis and the necrosis the pathway which are involved in the apoptosis so apoptosis can be produced or can be induced by the external factors and as well as the internal factors and then what is the relevance of the apoptosis in the overall biology of the organisms how it is actually affecting the other kinds of processes especially the development. Now the first question comes why the cell actually commits the suicide or I will say program death okay. So as we discussed right the cell is actually going to take up the nutrition and this nutrition is actually going to produce the energy in the form of the ATP or energy in the form of reducing equivalents such as NADH right and the purpose of these energy sources are that it is actually going to be used for the growth of the cell it is actually going to be used for reproduction and apart from that it is also going to be used for maintaining the cellular integrity okay. What is mean by the cellular integrity is that this energy is actually going to be utilized for maintaining the electrode potential or plasma membrane potential right. So actually what happened is that it is actually going to run the pumps right sodium potassium pumps or proton pumps and as a result it is actually going to throw the proton outside right and so that there will be a potential what is going to be

developed right.

So it is actually going to generate a potential across the plasma membrane because of that it is actually going to have the negative charge inside and positive charge outside or I will say it is actually going to have the negative polarity onto the plasma membrane and because of that if you have any object which is outside like for example the glucose molecule right. So glucose will not enter into the cell because it does not it is actually a charged membrane so this molecule cannot enter passively inside the cell it has to go through by a process of a receptor. So it actually has to go through with either through the receptor mediated endocytosis or it actually has to use the transporters to enter into the cell right and that's how it actually means going to decide or it is actually going to maintain an electrode potential and that electrode potential is helpful in terms of stopping the entry of the external factors. For example if you have a water molecule right the water molecule can easily enter into the cell right but it does not because the water is will actually going to use the water channels or water is actually going to be because if you can imagine that if the water will keep coming into the cell the cell will actually going to expand right so it is actually going to increase its size and ultimately it is actually going to burst right and that is what is going to happen when you are actually going to have the osmotic lysis. So these are the some of the things which actually cell opposes throughout its life right and for that only it is actually spending the energy in the form of the

ATP

or

NADH.

Apart from that it also require the energy for many more things actually it's okay. So once the cell is weak or cell is getting through a process of aging it actually has reduced its ability to produce the enough amount of energy and when it does not produce the enough amount of energy what it actually can do is it is actually has to cut down the activities it has to cut down the activities so it actually will cut down the active first thing what it will actually going to do is it will cut down the activity of cell disease right this means it will enter into the G0 phase just to conserve the energy right it will enter into the G0 phase. In the second event what it actually will do is it will actually going to you know it will not going to use the you know so if if the energy is still going to be on a lower side then it will decide whether I should be able to maintain my integrity or not because cell is cell only until it is actually having a integrate system actually okay. Once it actually become porous right and it can allow the cells allow the entry of the external molecules then the cell will eventually going to burst and since the cell will actually going to burst and release its content into the external media it is not going to be good for the organism because if you release the cellular content in one shot right it is actually going to create the you know the disturbance to the homostasis. So because to avoid this what it actually going to do is when it reaches to that point where it will not be able to maintain or it will not be able to generate the enough energy what it actually going to do

is it is actually going to induce the programmed cell death which means it was actually going to say that okay I am no longer be in a state that I can be able to maintain my integrity so let us go for the death pathway and then only it is actually going to go with the death and as I said you know death could be induced by the apoptosis or it could be necrosis.

So apoptosis or the cell this programmed cell death is a fundamental process in a multicellular organism which play a crucial role in the various biological processes such as development, tissue homostasis, immune response and the elimination of the damage or the potentially harmful cell. So this is you don't have to worry about all this because this is all we are going to discuss at the end of this lecture how the apoptosis is involved in development, tissue homostasis and elimination of the potentially harmful cells and all that. Apoptosis is a tightly regulated physiological process that rapidly removes damaged, mutated or virus infected cell within organisms and the major feature of an apoptosis that it is a controlled and the ordered process in contrast to the necrosis. Okay so apoptosis is also called as programmed cell death which means you are actually going to do the programming to have the death of that particular cell right that say it is actually going to be a controlled and ordered process because you are going to do a programming I'm sure many of you probably know about programming right so you actually gives the steps or you are going to give the command to the particular computer right okay go with this go with this go with this right and ultimately execute this and do this job right so the same way you are actually going to go with the stepwise instruction to the cell and eventually the last instruction would be that okay induce the death right then we have the specific signaling pathway and the molecular event which actually going to drive the process and this all how you are going to do you are actually going to do this by having the specific signaling pathways because these signaling pathways are actually going to bring the molecular players and that's how they are actually going to make the process more controlled then we cellular components are dismantled leading to the cell shrinkage then DNA fragmentation occurs resulting into the characteristic ladder pattern apoptotic bodies are apoptotic bodies small membrane bound vesicles which are actually going to be formed so shrinkage of the cell cinderella ring and as well as the formation of the apoptotic body these are the hallmarks of the apoptosis so if I have to identify an apoptosis in a cell I will see whether it is forming the DNA fragmentation or not if it whether there are apoptosomes are formed or not and whether the cytosol is shrinking or not right so that is are very different from the from the necrosis now what are the steps are involved in apoptosis so apoptosis actually is going to start from the cell right so its first step would be that the cell is actually going to receive a signal right this signal could be from the external signal or it could be an internal signal okay so it actually first step is that it actually going to receive the stimulus this is stimulus would be external stimulus or it could be an internal stimulus then once they receive the stimulus it actually going to

induce the shrinkage as well as the breakdown of the cytoskeleton which means it's actually going to make the cell more flexible so that it can actually be able to get condensed then we have the dense cytoplasm and the packaging of the organelles then we have the condensation of the chromatin this process is called as the pyknosis then we have the induction of the caspase activated dns activity and that will induce the fragmentation of the DNA means this is the fragmentation of the genomic DNA and breaking of the nucleus the step 6 it is actually going to form the blobs and then step 7 there will be a cell break and the formation of the apoptotic bodies so these are the formation of the apoptotic bodies this means the individual cellular content is actually going to be encircled into the plasma membrane and that is actually going to be apoptotic bodies these apoptotic bodies are not going to release the content they are actually going to be taken up by the macrophages and ultimately is that it is actually going to be phagocytose so that's will be that's why this is actually going to be a very safe way of removing the particular dead or damaged tissue without even causing any kind of adverse reactions now what is the difference between the between the apoptosis and necrosis so in the apoptosis you are going to have the intact cell membrane and membrane blubbing is going to occur right so this is actually going to happen right so it is actually going to maintain intact cell membrane but there will be a blubbing into the membrane okay compared to that in the necrosis there will be a disrupted cell membrane and there will be a loss of membrane integrity right so this there will be a loss of membrane integrity this means it is actually going to start losing the cellular content as soon as the necrosis is going to occur into the cell right and in the step 2 it's begin with the cell shrinkage and the condensation of the nucleus and then the picnosis occurs which means the condensation of the chromatid and followed by the karyohexesis which is the fragmentation of the nucleus and ultimately there will be a formation of the apoptotic bodies so these are the formation of apoptotic bodies which means it is actually not going to release the cellular content into the external media or the outside instead it is actually going to release the apoptotic bodies and these apoptotic bodies are actually going to be psychocytosed by the macrophages so that they will be very very very safely they are actually going to be removed from the environment compared to that the in the during the necrosis it is actually going to be it begins with the swelling of the cell right so initially the cell is actually going to swell because it cannot maintain the integrity right followed by the picnosis and the karyohexesis occurs followed by the karyolysis which means there will be a dissolution of the cytoplasm and ultimately there will be a complete lysis of the cell which will and there will be no apoptophome which is going to be formed and then ultimately it is actually going to release the complete cellular content and as a result of that it is actually going to cause a huge inflammatory reaction because this cellular content is going to be you know it actually going to attract the many type of cellular many type of immune cells and once they come to that they will actually going to try to clear this and they will actually going to cause the

inflammatory reaction which means they are actually going to secrete the inflammatory molecules and as a result there will be more damage into the vicinity where the necrosis is going to occur. Now coming back to the apoptosis there are multiple pathways which are involved into the apoptosis okay so in the apoptosis you have the three steps first is initiation second is execution the third is the psychocytosis okay because after the execution it is actually going to form the apoptosome and or apoptotic bodies which are and these apoptosomes are actually going to be psychocytosed okay so this two and step number two and three are actually going to be the same for both the pathway but the initiation pathway can be different for the two pathway okay because the first step is the stimulus right so if the stimulus is internal right then it is going to be intrinsic pathway if the stimulus is external then it is going to be intrinsic pathway or extrinsic pathway so that is the only difference okay if it is a internal factors such as starvation loss of nutrition generation of free radicals and all other kind of things then it is going to be induced the intrinsic pathway if it is the extrinsic pathway then it could be external factors such as the immunological molecules or the external ROS or drug molecules and so on okay so if you have these then it is actually going to induce that external factors it is going to cause an intrinsic pathway extrinsic pathway whereas for intrinsic pathway it could be the starvation it could be the ROS internal ROS or it could be any other kinds of anomalies that actually going to be responsible for the intrinsic pathway once the initiation is done then it is actually going to have the execution so in the execution you are going to have the caspases different types of caspases and then ultimately it is going to form the apoptosome and that actually is going to be the phagocytosed by the phagosomes so let's first discuss about the intrinsic pathway so in the mammals the signals that induce the apoptosis can either originate from the inside of the cell that is the intrinsic pathway or from the outside of cell which is called as the extrinsic pathway both signaling cascade ultimately leads to the caspase activation which in many define the number of in define the point of no return for the cell death these dead signaling at the events appear to be funneled to the mitochondria before the execution of the death by caspases into the mammalian cells so what are the molecular players involved into the apoproteins pathways first is caspases so caspases are the proteases are caspases are the group of protein involved into the apoproteins process they are called they are so called because they contain a key cysteine residue into the catalytic site and selectively cleave the protein at a site and selectively cleave protein at a site just C-terminal to the aspartate residue and caspases are the proteases all caspases are initially made as the pro caspases which means they are actually going to be produced as the inactive protease and ultimately they are actually going to be you know generate the active protease and by doing so they are actually going to be under the fine control which was cell will contain the inactive protease but it will actually going when it's going to get the signal that okay convert the inactive protease to the active protease it is actually going to be start the cascade of reactions then we also have the pro apoptotic factors and as well as the anti

apoptotic factors so within the pro apoptotic factor these factors are promoting the apoptosis such as bags bag with box bits back etc whereas we have the anti apoptotic factors these factors are inhibiting the apoptosis this means they are actually promoting the growth so these are called BCL 2 BCL-XL MCL 1a etc and the ratio between the anti apoptotic BCL 2 and the pro apoptotic batch protein determine whether a cell will actually going to live or it is actually going to die with by the process of apoptosis.

Now in the intrinsic pathway or the mitochondrial pathway you are going to have the you're going to have the initial signal right so there will be a signal which is like DNA lessons for example if you have generated the mutations or DNA damage and these DNA damage are beyond the repairing okay because some of the DNA damage can be could be repaired but some of the DNA damage could be so much that it will decide that okay it is not worthwhile to you know to repair the damage so instead of that it is actually going to in take up that as a initiation signal and then it is actually going to activate the serine and the atm cm kinase and that in turn is actually going to activate the production of the p53 transcription factors and the p53 is actually going to activate the downstream molecules like puma then puma is actually going to activate the BACs within the cytosol and the BACs activated BACs becomes the mitochondrial membrane bound and as a result it is actually going to open the voltage-gated channels right and it actually going to form the membrane attacking complex and once that happens it is actually going to release the cytochrome C from the mitochondria into the cytosol and that is actually going to be initial event right once the cytochrome C is actually going to be released it is actually going to form a complex with the APA one right and once they will form a complex to meet each other it is actually going to form the apoprosome which means the APA one and the cytochrome C when they come together they are actually going to form the apoprosome and apoprosome is actually going to activate or it's going to recruit the inactive caspase 9 and it is actually going to be activated to form the active caspase 9 so from the active caspase 9 it is actually going to act on to the pro caspase so it's actually going to form the pro caspase 3 and it's actually going to act on the pro caspase 3 okay and then it becomes activated caspase 3 so these are called execratory caspase these are called initiator caspase so activated caspase 3 is going to form an activated caspase 3 is actually going to activate the caspase activated the DNA's the once the caspase activated DNA is formed it will actually go inside the nucleus and then it is actually going to form the it is actually going to start chewing the DNA but it is very specific so it is actually going to chew the DNA after every 180 nucleotides okay and that's how it is actually going to form a ladder like things right because it's every ladder is actually going to be different from each other by a number of 180 so first DNA is going to be 180 first second DNA is 360 then 540 then 720 like that okay so that's why it is actually going to form a ladder like this okay where you are in between the two DNA is back bands it is actually going to have a difference of 180 base pairs okay this

all you will understand when we are going to talk about the DNA packaging so it will understand why there is a difference of 180 base pair right and apart from the caspase activated DNA's the caspase 3 is also going to be start acting on to the cytosolic as well as a nuclear protein and it is also going to start chewing up those proteins so as a result it is actually going to disturb the cellular machinery and at the end it is actually going to cause the cell death then we have the extrinsic pathways extrinsic pathway could be of two types extrinsic pathway where you have the TNF pathway or the fast like fast pathway in a TNF pathway you are actually going is going to be activated by the TNF alpha which is going to be secreted by the macrophages and other immune cells whereas fast pathway is actually going to be activated by the fast ligand and fast ligand is present on to the some of the immune systems or immune cells so first discuss about the TNF pathway so in a TNF pathway the acceptor TNR1 so this is actually going to be the receptor what is being responsible for the TNF pathway so on one side it is actually going to have the so this is an inactivated TNF receptor 1 and why it is in the inactive because it actually contains the death domain which is going to be captured right and on the top it does not have the TNF alpha right so it is receptor TNF R1 contain our intracellular part contains the death domain and this present on cell which is going to receive an apoptotic signal the death domain is silent prior to the apoptotic signal therefore it is called as inactivated TNF alpha TNF R1 the TNF R1 receives an external signal triggering molecule called TNF alpha cytokines which initiate apoptosis conformational changes occurs on to the intracellular part of death domain when the TNF alpha binds with the TNF alpha receptor 1 the death domain contains an inhibitory protein called SODD or the silence of death domain that keeps the death domain silent therefore the cell survives so what happened is that inactivated TNF alpha when it enters the TNF alpha will actually go into bind this TNF alpha can come from another micro immune cell such as macrophages and as soon as that happens that the domain is actually going to be active and then it is actually going to recruit the other cytosolic factors like thread FADD and then ultimately it is actually going to activate the caspase 8 ok so it's actually going to cleave the pro caspase 8 and that is actually going to make the active caspase and once the active caspase 8 is going to form it is actually going to induce a cell death by taking the help of the caspase 3 so after this actually it is actually going to activate the caspase 3 ok then we have the fast pathway right so in the fast pathway you have the two different types of protein one is signaling protein or I will say cytotoxic T lymphocyte for example right so example is cytotoxic lin lymphocytes where you are going to have the target protein target protein is the cell which actually going to be go through process of apoptosis right so this is the apoptotic cell ok and this protein the target protein is actually going to have the fast receptor whereas the signaling cell is actually going to have the fast ligand ok and once the fast ligand which is present on to the signaling cell is actually going to interact with the fast receptor what is present on the target cell it is actually going to induce the apoptosis and so how it happens you have the

fast ligand what is present on to the signaling what is present on to the signaling cell and then you also have the fast receptor what is present on to the target cell and when they interact with each other it is actually going to give you the signal and once they give the signal the death domain is actually going to recruit the downstream effector molecules and as a result it is actually going to activate the pro caspase 8 to form the active caspase 8 and the active caspase 8 is actually going to activate the pro caspase 3 to form the active caspase 3 and once the active caspase 3 is going to form it is going to act on to the genomic DNA and as well as the cytosolic protein and that's how it is actually going to induce a cell death so sickling cell is an immune cell which is called as cytotoxic telomophocytes the cytotoxic telomophocytes express a protein which is called as fast ligand the LIS ligand initiate the apoptosis through a series of reactions the fast ligand binds to the target cell through a fast receptor present on it binding of the fast ligand with fast receptor send the first apoptotic signal fast receptor contains the intracellular death domain on binding the dot domain recruit a FADD that the fats associated with death domain adapter molecule that comes and binds to the death domain of the fast receptor the death effector domain or DED of FADD molecule further recruit caspase 8 which gets activated and form the caspase 8 then a bunch of molecule existing together which are called as the activated fast receptor or the FADD adapter molecule DED and a caspase 8 enzyme form a single complex called the disk right or death inducing signaling complex the death caspase cascade starts when caspase 8 is released from the disk and what happened that caspase 8 is actually going to activate the caspase 3 into the caspase into the caspase active caspase 3 and the active caspase 3 is actually going to act on to the genomic DNA and as well as on to the cytosolic as well as the nuclear protein and eventually it is going to induce the cell death okay. Now what will be the what will be the relevance of these apoptotic pathway apart from the death okay it also has a relevance in many other features of the organisms. So the one of the major the area where the apoptosis has a relevance is the development right. So apoptosis is necessary in many developmental process during the limb formation the separate digits evolve by the death of the interdigital mesenchymal tissue. So what you see here is that this is the this is the two hands actually.

So in when you when the baby is within the womb right or whether during the developmental stage it actually has the hand which actually contains the membranes okay which actually contains the skin actually right and these hands are called raised hand okay and once baby is born these cells which are actually be a part of this web is actually going to be digested or which is actually going to be killed by a process which is called as programmed cell death and as a result we are actually going to have the individual heads. In the case of frog for example right frog does not have the individual fingers right it actually has the finger like this because it helps the frog to float onto the cell onto the water actually. Then we have the deletion of the cell no longer needed such

as the amphibian tadpole tail during the metamorphosis. So when the amphibian when the frog is developing from the tadpole so this is the tadpole right it contains a very long tail right but this tail is getting regressed when it is actually forming the adult frog actually what happens is this the cell what is present in this tail is actually going to be you know removed by a process which is called as the apoptosis. Then demise of the cell showing the structuring of the hollow tissue so this is this is what happening when you are actually going to form the from the body cellome right okay.

So during the development you are actually going to have the degradation of these cells or death of these cells so that's how it is actually going to form a tube like structure for example the development of the element chicken off. Then we have the formation of the reproductive organs and the massive cell death occur during early stage of nervous system greater than 50% of the cell neural cell actually are going to die. So what is the conclusion? In conclusion the apoptosis is a crucial process in a multicellular organism it is a controlled and ordered leaf form of cell death that play a vital role in development tissue processes and the immune system. Apoptosis is regulated by the specific pathway and involve the molecular event leading to the cell dismantling and the formation of apoptotic body. This regulation of apoptosis associated with the various type of issue one is one of such thing is formation of the you know formation of the fingers with the web actually.

So in some of the kids when they are born these webs are already present because the apoptosis does not occur and they are actually going to have this and in those cases what people do is they will actually going to be surgically be removed by the doctor actually or sometime what you see here is that the fingers are actually fused with each other. So in that case is also it is actually going to be surgically be removed because the apoptosis did not get induced in those particular people actually. Then understanding the mechanism and the significance of apoptosis opens venues for the septic intervention and shed light on the fundamental process of life and death. So this is all about the apoptosis. Now if you would like to study the apoptosis in a cell you can actually be able to use some of the classical features.

For example in there when you have a cell and if it is inducing if you are suppose treating this cell with a anti-cancer drug. So eventually what happened is you are actually inducing which you are so you are actually activating the in extensive pathway okay or you can actually inducing the intrinsic pathway whichever the you know because it depends on the different types of cell right. So different types of molecules so you can actually be able to have both of these. So what happened is the cell is actually going to show you the three important features which can be exploited for studying the apoptosis. One is there will be a shrinkage or cell shrinkage right.

The second is there will be a DNA damage and the third is there will be a membrane polarity and all of these methods apart from that you can also have the caspase activation. So you can actually be able to if you want to study the apoptosis you can use any either of these methods you can actually go with the DNA damage and look for the cell shrinkage you can look for the membrane polarity. Membrane polarity there will be a loss of membrane polarity there will be a loss of molecules from that particular cell. So one of the very easy thing is that you can actually be able to stain the cell with particular type of dye and the dye is actually going to show you whether there will be a cell damage membrane polarity or the caspase activation. So you can actually have the different types of substrate what you can actually use or you can use the dyes.

So one such approach is that where you are actually can use the combination of dyes which is called as a acridine ion propidium iodide and that you can actually be able to use for the monitoring the apoptosis. So what is the basic principle of this particular assay? Propidium iodide is a membrane impermeable dye that selectively bind the DNA by inter-calating into the double helix in live cell which has an intact membrane which means which actually has the intact membrane integrity. Propidium iodide is unable to enter the cell and therefore does not stain the nucleus. However in cell with the compromised membrane integrity such as dead cells, propidium iodide can penetrate the plasma membrane and bind the DNA resulting into the red fluorescence. So propidium iodide is actually going to give you the intense red fluorescence.

Okay to analyze apoptosis using acridine orange and propidium iodide a mixture of the dye is typically added to a cell suspension. In that cell suspension what happened is the live cell is actually going to appear green because the propidium sorry because the acridine orange is actually going to give you the green fluorescence when the cells are live but it going to give you it is going to give you the orange or red fluorescence when the cells are under the apoptosis. So this is for the live cell, this is for the apoptosis cell. So in a live so all the live cell is actually going to appear green. Apoptotic cells are going to appear orange or the red depending upon the amount of or the degree of fragmentations and the dead cells are actually going to appear as red right.

So these you are actually going to have three colors you are going to have green which is for the live, you're going to have the orange which is for the apoptotic cell and you are going to have the red which is for the dead cell right and you can easily be able to identify all of these colors in a technique which is called as flow cytometry and remember that we have very extensively discussed about the flow cytometry in our previous lecture so you can be able to utilize the flow cytometry for this particular type right. And these are the some of the material what you require for performing the AOPI

apoptosis assay and this is the protocol. So you have to take the 10,000 cells you are going to treat them with a substitutable anti-cancer compound or any other compound which you are thinking that it is actually going to induce the apoptotic cell death and then you can actually be able to follow this and it is actually going to give you the staining for the cells. So working consultation for the AO is 1 to 2 microgram whereas for the PI it is 20 to 50 microgram and it is actually going to add it to the cell 15 to 20 minutes before acquiring the data on a flow cytometer recuberance and as far as the data acquisition is concerned the after staining analyze the staining cell using a flow cytometer equipped with the appropriate filter for ocidine, orange and propidium iodide. Adjust the flow cytometer setting for the appropriate fluorescence and forward and side scatter parameters.

Run the stain cell sample on the flow cytometer collecting data for at least 50,000 events right. Analyze the acquired data using flow cytometer software plot the scatter plot with PI on the y-axis and the occludally orange on the x-axis. The first quadrant represent the healthy and the live cells and you can partition that plot into the four quadrant the second third and fourth quadrant represent the early apoptotic, late apoptotic and dead cell respectively and you can do the data analysis with the help of a software which is called FCS-5. You can use any other software this software is not exclusive okay and ultimately what you are going to see you're going to see the results okay. So, what you are going to see is this is the control right and this is the treat example okay.

But before discussing about the results we can actually take you to my lab and where the students are actually going to show you the complete protocol and how you can be able to analyze or do the acquisition of these data into the flow cytometer and how you can be able to analyze the data to get this result. Hello everyone, in this video we will be discussing about how to perform live dead cell staining using adiridin-orange and propridium iodide on FaxA. So the basic principle is that the adiridin-orange is permeable to both the live and dead cells whereas the propridium iodide is only permeable to late apoptotic and necrotic cells. So this property of adiridin-orange and propridium iodide lets us to recognize what population of cells are in late apoptotic or early apoptotic or necrotic cells. So coming to the procedure the first thing we do is that we treat the cells from the 100 mm cell culture dish and then plate 1 million each in the untreated and the treated well.

So after 2 to 14 hours of adherence we treat the samples according to our requirement and then let's say for that we are treating for 24 to 48 hours then after the appropriate time we crystallize the cells collect the pellet wash it 2 times with PBS and then resuspend the pellet in 2% beta 2 1 serum in phosphate percolate. So after we have resuspended the pellet in 2% and in phosphate percolate we give the appropriate

adridin-orange and propidium iodide treatment. The working concentration for adridin-orange is 0.5 to 1 microgram per ml whereas the working concentration for propidium iodide is 1 to 5 microgram per ml. So we add the dyes just before taking the data or we can just give 10 to 15 minutes of incubation for the dyes to bind to the cells so that and after that we record the data on the fat cell content.

So after adding the adridin-orange and propidium iodide to the cells we have to acquire the data on the CellQuest Pro software. So the first thing we do is open the CellQuest Pro and connect it to cytometer and then we need the counters the detector and amps and the status. For the adridin-orange, propidium iodide staining we need two dot plots. One is for the FSC, SSC for the forward scattering and the side scattering and the other one is for the FL1 and FL3. So the FL1 plot is on the x-axis whereas the FL3 plot is on the y-axis.

The FL1 plot is for the acridin-orange and FL3 plot is for the propidium iodide. After taking the plots we have to set the directory and save the data in our required location. In the detector and amps we have to remember that we have to set the population of the healthy cells in the first quadrant that is 10 to the power of 1 and 10 to the power of 1. So after we set the untreated cells in the first quadrant and then we analyze the treated cells and then we can say whether there is any shift in the fluorescence in the untreated and the treated cells. For the treated cells in the third and the fourth quadrant that represents the apoptotic and the necrotic cells.

So now we will be taking the sample but before analyzing the data we have to set the number of events to the 5000 and then keep it on setup and first we will see whether the events are coming properly or not. Now we press acquire. As we can see that in the FSC and SSC plot we can see the events coming near 0 0 that represents the healthy population as well as in the FL1 and FL3 we have set the healthy population between the 10 to the power of 1 and 10 to the power of 1. So this represents the first quadrant. We will show in detail how to do the quadrant analysis in the FCSX-5 software.

Now that we have set the population in the first quadrant we will remove it from setup and acquire the data. After the untreated samples we have to take the treated samples on the same parameter description which we have set for the untreated cells. Now we change the sample in the sample injection port to the treated sample. We have to remember that we don't have to change the parameters or else we will never be able to say whether there is any shift in the untreated or the treated cells if we change the parameters. After changing the sample now we have to choose the directory for the treated cells and then change the name also to treated and also the file count to 1.

Then press OK and then now we acquire the data on the same parameters. As we can see that there is some shift in the population of the cells. The population is having a little bit more fluorescence than untreated cells which represents the apoptotic and necrotic cells in the third and the fourth quadrant. In the fourth quadrant mostly the necrotic and the dead cells are present whereas in the third quadrant the late apoptotic cells are present. After we take the untreated and the treated samples after we acquired the data for the untreated and the treated samples we have to analyze the data in the FCS5XPress Pro software using quadrant analysis.

In the quadrant analysis we can see how much populations of cells are present in which quadrant and therefore we can identify the number of healthy populations and the apoptotic and the necrotic cells. So after acquiring the data in the FACTS equipment we have to now analyze the data in the FCS5XPress software. So the first thing we do is we open the new layout and change the orientation to landscape and then now we input the data and then first thing we do is we take the untreated file and then open the dot dot. We need two dot plots the one is the FSC SSC and the other one is the FL1 FL3. The FL1 FL3 dot dot shows the live and dead cells staining.

The FL1 is responsible for the acridin orange whereas FL3 is responsible for the propidium iodide and now we take the treated file and again we select the dot plot. In dot plot as well we need the FSC and the FL1 FL3 plot. As we can see that there is a difference between the untreated and the treated sample. Now we have to find out how much percentage of the cells have gone actually the apoptotic or necrotic. Like we have to divide the population of cells into four quadrants using the quadrants option.

So we go to the gating and then take the quadrant and then apply it on the FL1 FL3 plot. We have to apply it in such a way that we cover all the cells in the untreated plot. So let's say that in the untreated plot we are having 93 92 percent on live cells and 7 percent are in the early apoptotic phase and then after applying the quadrant to the untreated we have to apply the same quadrant to the treated one in order to find out the difference between the two. So just we click on the quadrant and then we copy and paste on the treated one. So in this way we can say that in the first quadrant in the untreated sample we have approximately 93 percent of live cells whereas in the treated one we have like in the treated one we have only 35.

7 percent healthy whereas the 33 percent have gone are in the late apoptotic and 29 percent are necrotic cells. So in this way we can use acridine orange and propidium iodide to determine the healthy the apoptotic and the necrotic cells built in different treatments and also we can establish a relation between different concentrations of treatment and the number of live and dead cells in any experiment. So this is the way we

analyze the we analyze and process the data on the FACTS equipment in order to do the live and dead cell training. Hopefully this video was helpful. So once you analyze the data you are actually going to get this data right you are going to get these two curve right or these two plots and these are called as the dot plots okay and where you are going to have the the checkerboard analysis.

So this is called as checkerboard analysis right and in the checkerboard analysis what you're going to do is you're going to make a checkerboard in such a way that you are going to keep all the healthy cells in the first quadrant. So this is a quadrant 1, this is the quadrant 2, this is the quadrant 3 and this is the quadrant 4 okay. In the quadrant 1 you are going to have the low fluorescence for the FL3 and low fluorescence for the FL1 which means it is actually going to be the healthy cells because they are not taking up the strong fluorescence signal or they are not taking up the dye from the they are not taking up the dye which means they are their membrane potential, membrane polarity is very high and that's why they are not allowing the dye to enter right. When you treat these cells the cells are actually going to be apoptotic and as well as the dead so they will actually enter into the next quadrants. So for example this is the quadrant number 1 which is the healthy cells right this is the quadrant number 2 which is actually the early apoptotic cell which means now just the DNA damage started actually so and that's why what you see here is that it has a very high signal for FL1 but it has a very low signal for FL3 right this means these are early apoptotic and this is the late apoptotic because now the DNA is compromised and so it actually has a very high signal for FL1 and also has a high signal for FL3 and that's why these are the late apoptotic cells and this is the cell where you have the low fluorescence like low FL1 and high FL3 and these are the dead cells right.

So what you see here is that in the treatment you have the 27% dead cells 36% late apoptotic cells and the 11.8% early apoptotic cells and whereas 23.5% healthy cells. So these are all about the apoptosis. Now when you are doing these kind of essays you always have to take a lot of precautions when you are analyzing the data when you are making a checkerboard and all that so that you should not make mistakes so that it should be make you biased or you should not be able to get the you know the data which is unreliable.

So these are the some of the precautions what you have to follow you have to acquire the samples under the 4 degree and you also have to require other kinds of treatments other kinds of precautions what you are supposed to take. So this is all about the apoptosis or the programmed cell death what we have discussed we have discussed about why the cell is entering into the apoptosis what are the critical factors which are inducing the apoptosis and what are the different types of pathways which are responsible for

apoptosis. So we have discussed about the intrinsic pathway and we have also discussed about the extrinsic pathway within the intrinsic pathway we have discussed about the molecular players which are governing the intrinsic or the extrinsic pathways so and at the end we have also discussed about how you can be able to study the apoptosis with the help of the flow cytometry. So with this I would like to conclude my lecture here in a subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you. you

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 03
Basics of Biomolecules
Lecture-11 Biomolecules (Part 1: DNA)

Hello everyone! This is Dr. Vishal Trivedi from the Department of Biosciences and Bioengineering IIT, Guwahati. And what we were discussing? We were discussing about the different aspects of the molecular biology. And so far what we have discussed? We have discussed about the basic properties of a biological system. In this context, we have discussed about the structure of the biological cells. So we have discussed about the structure of the Bocaryotic cell and we have discussed about the structure of Eukaryotic cell.

While we were discussing about the structure of Eukaryotic cell, we have also discussed about the structure of the different organelles and their functions and their contribution into running the different types of functions or the activities within the cell. In addition to that, in our previous lecture, we have also discussed about the cellular metabolism and how the cell is producing the energy with the help of the different types of catabolic reactions and anabolic reactions. So we have discussed about the anabolic reactions also which are responsible for production of the different types of biomolecules or the proteins. And then we also discussed about how the cell is actually dividing and how what are the different stages it is actually going through.

So we have discussed about the cell cycle and how the cell cycle is tightly being regulated. While we were discussing about the cell cycle, we have also discussed about the some of the technical and as well as the experimental aspects, how you can be able to study the cell cycle with the help of the flow psychometry. In addition to that, we have also discussed about how you can be able to study the mitosis and meiosis with the help of the different types of microscopy techniques. And then at the end, we have also discussed about the apoptosis and cell death. So and there also we have discussed about the different types of pathways.

So we have discussed about the intrinsic pathway and as well as the extrinsic pathway, what are different molecular players which are actually participating into these type of different pathways. And we have also seen how we can actually be able to study these events in the DNA cell with the help of the flow psychometry. So now it is clear that there are so many functions which are happening or there are so many activities what is happening within the cell. And all these activities, if you summarize, you will summarize

that these activities can be classified into three different types of activities. And all these activities can be sum up within the life related activities.

So you can have the life related activities. So first is the production of energy which means the catabolic reactions. We have discussed about the carbohydrate metabolism. And we have also discussed about the lipid metabolism and how the glycolysis and crepe cycle is contributing into the production of energy and how the beta oxidation is taking place in the lipid and how it is actually producing the stile coA and then the stile coA is getting into the crepe cycle and producing the energy. The second aspect is about the growth.

And as you remember that when we were using the energy that energy can be used for producing the growth with the help of running the anabolic reactions. And the third aspect is about reproduction. So growth is different from reproduction where the cell is actually going to perform the different types of divisions. And that is how the single cell is actually going to multiply and becomes double two cell, four cell, eight cell like that. Now what you see here is that if you want to go with the catabolic reactions or the anabolic reactions or the reproduction of the reproduction where you are going to do a division.

In this also there will be anabolic reaction what is going to involve in the production of different types of biomolecules. So all these activities are completely being regulated by the biomolecules at the molecular levels. And in this particular module, we are actually going to go through with some of the basic properties of these biomolecules. So what are the different biomolecules we are going to cover in this particular modules. So we are going to discuss about the DNA or DRC ribonucleic acid.

Why it is important to study the DNA because DNA is a source of the genetic information. This is anyway we are going to discuss when we discuss about the genomic DNA and other kinds of aspects of molecular biology. But for time being you can imagine that the DNA is actually the genetic material. Because when we were discuss about the genetic material, you will understand that that time we are going to discuss about the different types of experiments. Which are actually going to give you the idea that why the DNA is a genetic material.

Then DNA will go through with the activity which is called as replications. And with the help of the replications, one DNA copy will actually go into form the two DNA copies. And remember that when we will discuss about the replications, we will discuss about the replications in the poropyratic system and as well as replication in the eukaryotic systems. And then DNA is also going to participate into the activities which are called as transcription. So, from DNA it is actually going to generate the RNA and RNA is going

to have the same amount of information what the DNA is.

And this process is known as the transcription which is actually going to perform by the RNA polymerase. DNA to DNA is actually going to be catalyzed by the DNA polymerase. Now, once from DNA, you are actually going to synthesize the RNA. RNA or the ribonucic acid, it is actually going to participate into translation. Translation which is actually going to be responsible for production of protein.

So, RNA alone is not responsible. You are going to have the help of the other protein molecules to form the ribosomes and other machinery and then ultimately it is actually going to produce the protein. From the RNA, once you produce the protein and protein is very important because it is a building block for most of the biological system. For example, we have the different types of proteins. So, we can have the protein which is building blocks like collagen.

We can have the proteins as building block like actin, myosin which is actually being responsible for production generation of muscles. And because of these actin and myosin fibers, you can be able to walk and you can be able to run. So, it is not only that it is actually going to be a building block. For example, the collagen is actually the main fiber responsible for production or for the generation of the different types of bones because collagen is going to be calcified. And that is how it is actually going to produce the different types of bones.

Hairs, for example, the keratin and other things. Now apart from that, the protein can also be an enzyme. So, enzyme is actually I am keeping in a different category. So, enzymes are within the protein, they can also be an enzyme. And what is the function of the enzyme? Enzyme is actually going to participate in running the metabolism of the organisms.

What is mean by the metabolism is that it is actually going to run the catabolic reactions and the anabolic reactions. All these anabolic and catabolic reactions are actually going to be catalyzed by the different types of enzyme. Apart from that and what is the purpose of these? It is actually going to be responsible for the energy production. Apart from that, the enzymes are very actively being participate into the three major processes what are happening in the cell. What are these processes? These processes are called as the replication followed by transcription followed by translation.

Now all these are central pathways are actually being governed by the enzyme. Apart from that, the enzymes are also being playing crucial role into the molecular cloning. So, in this particular module, when we talk about the enzyme, we are actually going to discuss about the basic properties of the enzyme. And then we are not going to talk about the

enzyme kinetics and other kinds of aspects of enzyme. Because that you can easily be able to go through or that you can actually be able to study in other kinds of MOOCs courses.

Like for example, I also have another MOOCs course called enzyme science and technology. And if you want to study that part, then you can actually be able to study using this particular module MOOCs course. So, there is a MOOCs course called enzyme science and technology, which deals in detail about the enzyme related kinetics, how you can be able to solve the structures and blah, blah, blah like that. So, that we are not covering in this particular module. What we are covering is what basic properties of the enzyme.

And then we are going to talk about the specific properties of the enzyme which are going to play crucial role in the molecular biology related activities like replication, transcription and translation, and enzymes which are part of the molecular cloning. So, let us first start our discussion about DNA. So, DNA is a nucleic acid, which is a deoxyribonucleic acid. So, it is a deoxyribonucleic acid. You can have the two different types of nucleic acids.

So, and as I said, DNA is the major molecule which is responsible for carrying the genetic information from the one generation to another generation. And this all you are actually going to learn when we are going to discuss about the genomic DNA and genetic information, genetic material actually. So, most of the organisms whether it is a prokaryotic organisms or a eukaryotic organism has DNA as a genetic material, whereas minor fractions such as some of the viruses like the coronavirus or other kinds of HIV and all that has RNA as a genetic material. So, mostly the nucleic acid is going to serve as the genetic material. So, whether it is the DNA or RNA.

DNA or RNA is a biopolymer and it is acidic in nature. And that is why the nucleic acid is acidic in nature. In eukaryotic cell, animal or the plant, the nucleic acid is present within the nucleus, whereas in prokaryotic cell, it is present as a free form into the cytosol. And that is why we have discussed about the differences between the prokaryotic cell and the eukaryotic cell. And one of the major differences that the nucleic acid which is present within the nucleus in the case of eukaryotes, whereas it is going to be present outside as a free form into the cytosol.

The first nucleic acid was isolated by a scientist known as Vergic Mestre in the year 1868. And since then, we will actually discovering the new and new properties of this particular molecule. Now, what is the composition of the DNA? So, the nucleic acid is composed of the three components. You can have the phosphoric acid, you can have the base and you can also have the sugar. The phosphoric acid provides the backbone to the polymer, whereas the sugar work as the anchoring point for the nitrogenous bases.

The nine-membered nitrogenous bases give the diversity in the sequence of the nucleic acid. So, these are the three different components. You can have the sugar, right. So, you can have the two different types of sugar. So, we can have the ribose in the case of RNA, right.

So, we can have the ribose in the case of RNA or we can have the two deoxyribose in the case of DNA, because these are the, and remember that there is only one difference that 2 prime OH is present in RNA, whereas that 2 prime OH is missing in the case of DNA. And then you can also have the phosphate, right. So, the phosphate is actually working as a backbone, right. So, phosphate is running throughout, right. And in this phosphate, you are actually having the, you know, the sugar molecule which is attached to this.

And that sugar molecule is working as an anchorage point for the different types of nitrogenous bases. So, you can have the nitrogenous bases which are belonging to the pyrimidine or you can have the nitrogenous bases which are belonging to the purines. So, forget talking about the components, so phosphate backbone. So, phosphoric acid serves as the backbone of the molecule. So, you can have the two different chains, one is phosphate backbone on this side, the other is having the phosphate backbone on this side.

Then you can also have the sugar. So, you can have the five membered sugar, either it can be a ribose or the deoxyribose. So, in the DNA, you can have the deoxyribose where the 2 prime OH is missing. So, the five membered cyclic reducing sugar is present in the nucleic acid. These are the two different variants, the sugar molecule which contains the hydroxyl group at 3 prime as known as ribose, whereas it is deoxyribose if it is absent, okay. Based on the sugar, the nucleic acid is classified either the RNA or the DNA.

Ribose sugar is present in RNA whereas deoxyribose sugar is present in DNA. The purpose of sugar in the nucleic acid is to provide the attachment point for the nitrogenous bases. So, this is the sugar and it is actually going to have the attachment point so that you can have the nitrogenous bases attached to sugar. So, this is the phosphate backbone on this you are going to have the sugar molecule and on this sugar molecule, you are going to have the base, right. And that is how it is going to be have the same way, it is going to be on this side.

And that is how they actually interact with each other. Is anyway we are going to discuss in details when we discuss about the base pairing and other kinds of things. And then the nitrogenous base, so there are two variants. There are nine-membered conjugated double bond system, right.

And there are two variants. One is purine, the other one is called as pyrimidine. So, purine

such as adenine and guanine, the six-membered single-ringer system which is called as pyrimidine. So, these are called thymine, uracil and the cyclosilane. So, this is the adenine which is a nine-membered ring. And then you also have the another nine-membered ring which is called as guanine.

So, these are the nine-membered ring or I will say two-ring base, right. So, if you have two rings, it is going to be purine. If it is a single ring, then it is going to be pyrimidine, okay. So, pyrimidine is going to be either the thymine, cytosine or the uracil, okay. Whereas for the purines, it is going to be two-ring and pyrimidine, it is going to be one-ring, okay.

And the presence of nitrogenous base in DNA, RNA is predetermined. For example, the DNA has adenine, guanine, thymine and cytosine. And it does not contain the serosol, okay. Whereas the RNA has adenine, guanine, uracil and cytosine and strictly no thymine, okay. Now, the question comes, if you have the adenine, thymine, guanine and cytosine, why there is a strict or what is the base pairing? Which one will actually going to make the base pair with others and so on.

So, what is the rule about having the base pairing between these nitrogenous bases, right. So, the DNA is double-stranded where RNA is single-stranded, right. In most of the cases, RNA can be double-stranded in some of the plant viruses and some of the other special organisms who are actually going to have the RNA double-stranded, but mostly the double DNA is actually going to be double-stranded. The individual monomer responsible for making DNA or RNA is nucleotide and as a result, the DNA or RNA can be considered as poly nucleotide molecules.

So, they are polymers actually. Similarly, just like you know, the sugar is the polymer of the glucose molecule. Similarly, you can actually have the DNA which is a polymer of the nucleotides and that is why it is called as the poly nucleotides. In a poly nucleotide, you can and what is there in the individual nucleotides. So, individual nucleotide is having a nucleoside attached to one or more phosphate group and can be termed as the, if it is attached to the one phosphate group, then it is called as monophosphate nucleoside. If it is attached to the two phosphate group, then it is called as diphosphate nucleoside and if it is conducted to the three phosphate group, then it is called as triphosphate nucleotides.

This is what it is actually going to show. So, if the sugar is attached to the base, if the sugar is attached to the base, then it is called as nucleoside. So, if the sugar is plus base, then it is called as nucleoside. Now, if nucleoside is attached to the one molecule of phosphate, then it is going to be called as nucleoside monophosphate or I will say nucleoside. Then if it is attached to the two phosphate groups, then it is going to be called

as nucleoside diphosphate. Or if it is attached to the three phosphate group, then it is going to be called as nucleoside triphosphate.

And the base is attached to the sugar molecule with the help of the glycosidic bond. If the two prime OH is present, then it is going to be called as ribose. If the two prime OH is absent, then it is going to be called as DRC ribose. Each nucleoside is composed of the nitrogenous base attached to the sugar to the glycosidic bonds. So, when the nucleoside is going to be attached to the phosphate, then it is also going to be called as nucleotide.

So, when the attached with phosphate, then it is actually going to be called as nucleotide. So, nucleotide monophosphate, nucleoside diphosphate, nucleotide triphosphate. Now, nucleotide has the free hydroxyl group at the three prime carbon and a phosphate group at the five prime of the sugar. This is what is going to show. So, it has actually the five prime phosphate and actually going to have the three prime OH.

And that actually provide some kind of the orientation or polarity into the molecule. So, the first nucleotide has the free phosphate group and the three hydroxyl group which are going to make the bond with the phosphate group at the five prime of the next nucleotide. So, this one is actually going to make a bond with the phosphate group by the subsequent nucleotides. And that will continue. So, that is why you are actually going to have the five prime free phosphate group on one end.

And when it ends, then you are actually going to have the three prime free OH groups at the other end. And that is why it actually provides the polarity into the molecule. So, the propagation of the nucleotide along the length of the chain give rise the poly nucleotide. As a result of each poly nucleotide chain has the free five prime phosphate group on first nucleotide and the free three prime hydroxyl group on the last nucleotide. So, this is the first nucleotide on which you are actually going to have the five prime free phosphate group because it is not attached to anybody.

So, this phosphate group is free. And on the last phosphate nucleotide, this end, this phosphate, the sugar what is attached, it is actually going to have the free phosphate OH group on the three prime OH. It gives the polarity to the poly nucleotide chain and it runs in the direction from the five prime to the three prime. So, because of the simplicity and to make the things more systematic, we actually call this as running from the five prime to the three prime. Because five times you are actually going to have the phosphate group and on the three prime you are actually going to have the OH. Because so basically if you want to extend this growth, you are actually going to extend on this side not on this side because this side it is already been blocked by the five prime phosphate group.

So, DNA is a double standard where RNA is a single standard. This is there are exceptions that RNA could be double standard like some of the plant viruses and the animal viruses. Both strands of the DNA are held together by the hydrogen bonding between the bases attached to the sugar. So, you can have the A, you can have the G, you can have the T, you can have the C and all these are actually having the hydrogen bonding between them. So, adenine of one chain is always making the two hydrogen bonding with the thymine of the next chain. So, on one side if you have the adenine and the other side if you have the thymine, then it is actually going to make the hydrogen bonding of the two hydrogen bonding.

Whereas in the case of the same way if you have the guanine on this side, you can actually have the cytosine on this side. So, it is actually going to make the three hydrogen bonding. So, similarly the guanine of the one chain is making the three hydrogen bonding with the cytosine of the next chain. Now, the question comes why the adenine is making a pair with thymine and why the guanine is making a pair with cytosine. What is unique about this base pairing? It is possible that adenine can make a pair with cytosine and guanine is making a pair with thymine.

But that does not happen because of that there is a strict base pairing that adenine is always making a pair with thymine and guanine is always making a pair with cytosine. So, the question comes why there is such a strict base pairing and that strict base pairing the answer to this strict base pairing comes from their structure itself. So, you know that the adenine is adenine and guanine are the two ring structure and the thymine and cytosine is one ring structure. This means they are small, these are big. Now, you actually so and apart from that the groups what are attached to the adenine and guanine are actually different than the group what is present on the thymine and cytosine.

And you know for hydrogen bonding, it is these kind of stoichiometry and the distances are very important. So, why there is a base pairing of such a strict base pairing that adenine is going to make a pair with thymine and guanine is always making a pair with cytosine. So, question lies within their structure and as well as the groups what are present on to their rings actually. So, adenine or guanine is a purine and it is a 9 membered ring, 9 membered means it is actually having the 2 rings whereas, the thymine or cytosine is a pyrimidine which is a 6 membered ring.

So, this means a 1 ring structure. So, presence of both purines such means which means adenine and guanine right which are bulky actually. So, there are 2 rings right. So, they are actually bulky. So, if you put a peak adenine on one side and guanine on other side, it is actually going to have the steric hindrance. It is actually going to have the not enough

space between the DNA strands right and because of that they will be actually going to be too wide actually.

So, they are actually going to be too wide to get accommodated within the DNA structures right because the distances between the 2 strand of the DNA is going to be very strict right. So, they are not going to vary. In comparison to that for example, if you have 2 pyrimidines for example, the cytosine and thymine, then the cytosine and thymine are very small. So, they will not be placed very far away and then they will not actually going to have any kind of hydrogen bonding. So, hydrogen bonding is not going to be possible if the pyrimidine in pyrimidine is present because the pyrimidine is actually going to be with small because it is a single ring structure.

So, they will not be able to interact with each other and because of that the DNA structure is going to be unstable at this particular region right. So, if they will not be able to interact, this region can be broken very easily and it is actually going to form the loops and other kinds of bulbs right. The only combination what is possible is if you have the purine on one side, then you should have the pyrimidine on the other side, which means if you have the adenine on this side, then you can actually have the thymine on this side. So, in that case, the distances are also fine. So, if you have this combination, it is actually going to manage the distance because then the distance between the groups are going to be consistent.

Whether you have the bulky on this side and small one on this side or you have the bulky on this side and the small one on this side does not matter. The distance is going to remain the same. So, that is it is not going to distort, it is not going to destroy the DNA structure. Apart from that, the groups what are going to be present on to the purine is actually going to place the groups in such a way that they are actually going to interact and they are going to form the hydrogen bonding with the groups what is present on to the pyrimidines. So, this way, the most possible combination is that the purine on one side and the pyrimidine on the other side.

This means the amount of purine could be or is actually equivalent to the amount of pyrimidine. This way anyway we are going to discuss in detail. So, if you have the two purines, it is difficult because they are bulky. If you have two pyrimidines, they are small.

So, it will be too small to form the hydrogen bonding. And so the only combination would be that if you have the purine on one side and the pyrimidine on the other side and then it is actually going to have the perfect match for hydrogen acceptor and the donor side, which are present on to the purine and pyrimidine. So, that is why there is a strict base pairing, two chains and that is why the two chains are called as complementary to each other, which

means if you know the sequence on one chain, you can actually be able to predict the structure on the second chain. And that is why they are actually complementary to each other. Now, the question comes, what is complementary mean to you? What is complementary? So, it means that if I will provide you the sequence of nucleotide on the one strand, if I provide you this information about the nucleotide sequence on the strand one, it will let you to predict very precisely the sequence of the nucleotide on the other hand, because A, wherever you have A, it is actually going to interact with T and wherever you have G, it is actually going to interact with C.

Let us take this with an example. So, for every appearance of A, you will actually going to give the T and for every appearance of G, you are going to take the C. Let us take an example. For example, this is the strand one on which you are actually know the sequence, right. And if you want to know the sequence of the complementary strand, so this is going to be the strand number two. Remember that the strands are not only going to be complementary in terms of the sequence, they are also going to be complementary in terms of the polarity.

For example, strand one is running in the direction of 5 prime to 3 prime. This means it is running in this direction. This means the complementary strand should run in this direction. This means the 5 prime is going to be on this side and it will go run in this direction. This kind of information is very important and this aspect and concept is very important for you to understand because it is actually going to be used extensively when we are actually going to discuss about the replication and transcription because that time this complementary information and the concept of complementary is very, very important to understand.

So, let us see. So, strand one, you have the adenine, you have thymine, you have guanine, guanine, cytosine, cytosine. So, the first nucleotide, if you go to the first nucleotide, it is adenine. So, what I will do is I will put the thymine.

If it is thymine, then I will put the adenine. If it is guanine, I will put the cytosine. If it is guanine, I will put cytosine. Cytosine, I will put guanine like that. And that is how you can be able to have the generation of the second strand. And since it is starting from the 5 prime, I will put the 3 prime because as I said, the strands are complementary to each other not only in terms of sequence, but also in terms of the polarity. So, the individual strand of the DNA runs in the direction of 5 prime to 3 prime and on the other strand runs in the direction of 3 prime to 5 prime.

Hence, both strands are running in an anti-parallel direction to maintain the base complementarity. The presence of complementarity in base pairing and running of strand

in the anti-parallel direction allows the precise duplication of DNA through a process known as replication. And this is all we are actually going to discuss in detail when we are going to talk about the replications. So, remember that this is the 5 prime of sugar and this is

the strand 1.

And all these are the nucleotides. So, this is the base what is attached and then it is actually interacting with the, so this is adenine interacting with the thymine and this side this is strand 1. So, this is the strand 1 and this is running in from in this direction and whereas this is a strand 2 where it is running in the opposite direction. And remember that if you have adenine on this side, you are going to have thymine on this side. Now, let us talk about some of the rules what is related to this complementarity. So, understanding the base pairing requirement, the Chargaf has proposed a rule about composition of DNA.

Summary of this rule is as follows. Point number 1, the purines and pyrimidines are always going to be in an equal quantity, which means the amount of purines is going to be equal to the amount of pyrimidines. And this is understandable because whatever the amount of purine is going to be present on strand number 1, it is actually going to be present on the same amount of pyrimidine on the other strand. And that is why if you take the composition of the total DNA, the amount of purine is going to be equal to the pyrimidine. Because A is making a pair with T and G is making a pair with C. The amount of adenine is equal to the thymine and the amount of cytosine is equal to the guanine, which means A is going to be equal to the T and G is going to be equal to the C.

Not only that, the base ratio, which means A plus T divided by G plus C may vary from one species to another, but it will remain constant for a given species. And this is a very, very important information because if you calculate the AT by GC ratio, you can be able to say very precisely what is the species of, you can actually be able to identify the species of that particular organism. Because as I said, it will vary from one species to another species, but it will remain constant for a given species. So, he proposed that these ratio can be used to identify these species and you can actually be able to use this information to classify them.

Now, number 4, the deoxyribose sugar and the phosphate component occurs in the equal proportions. Now, the question comes if the DNA is double stranded, how it can be denatured to access the information of the nucleotide sequence. So, DNA double helix can be broken open if it is exposed to the high temperature or titrate with the acid or alkaline. Remember that the DNA strands are attached with each other with the help of the hydrogen bonding. So, this hydrogen bonding can be broken by two things, either you add something which is more polar, such as you change the pH or you add salt.

If you add the salt, the salt will actually interact with the base pairs and it is actually going to break. Other point is if you heat, if you increase the temperature, heat is actually going to break the hydrogen bonding between the bases. During this process, the hydrogen bonding between the two strand breaks and this process is known as the melting or the denaturation of the DNA. When the denatured DNA is incubated at low temperature, the separated strands re-associate to form the duplex DNA. This process is known as the re-naturation and this is very important concept to understand that when you are going to heat in the DNA, the two strands are actually going to be get depart, right, because the hydrogen bonding between the bases is actually going to be broken. So, they will actually going to get separated and when you are actually going to lower down the temperature, it is actually going to re-naturate, right.

And this concept is very effectively being used when you are talking about the technique which is called as polymerase chain reaction. So, the denaturation or re-naturation kinetics is used to understand the complexity of DNA and it has a wide application in amplifying the strand using a technique which is called as polymerase chain reactions. So, DNA denaturation and stability. So, if you do that, what you are going to see here is that it is actually going to give you the fraction of DNA what is present as the double standard versus single standard.

And if you plot this denaturation curve, you are going to get a sigmoidal curve like this. And this actually is actually going to give you the information when the 50% DNA is actually being denatured, which means when the 50% DNA is present in the double standard form versus single standard form. So, that is actually going to give you the TM of that particular DNA and that TM of that particular DNA is actually going to be a very, very characteristic to that particular species. It varies between the species and it also varies, where if the DNA is more complex or the, so it varies in terms of when the complexity of the DNA will actually go up, right. And that is why this particular type of denaturation curve can be used to understand the complexity of DNA without even going through the process of sequencing.

Now, let us talk about how you can be able to isolate the genomic DNA from the cell, right. Because you are actually going to use this information if you want to perform these kinds of experiments like where you are actually going to understand the complexity of DNA and all other kinds of things. So, what you are going to do is you are going to, so we are not getting into the detail of the protocol, like how you are going to put the different types of reagents and all that. What you are going to do is first you are going to lyse the cell with the different detergents containing so that it will actually going to prepare the, lyse the cells. So, you are going to lyse the cells and once you have the lysate, right, so it is actually going to contain DNA, it is actually going to contain the DNA and also going

to contain the protein. And it is also going to contain the minor quantity of lipids, right, because it is going to have the lipids from the plasma membrane.

So, these are the three biomolecules what is going to be present in this particular lysate, right. Then what you are going to do is, and DNA is actually going to be present in complex with protein because you know that the DNA is always making a pair, making a complex with protein because DNA is negatively charged. So, it binds the positively charged histones and that is how it is actually going to be packed within the nucleus. This all we are going to discuss in detail when we talk about or when we discuss about the genomic DNA in detail, right, when we are going to talk about the genetic material. So, then the second step is you are going to do enzymatic digestion. So, you are going to treat the cells with the digestion buffer, right, and the digestion buffer is actually going to contain the protease which is called as protease K and the SDS and it is actually going to release the genomic DNA from the DNA protein complex.

Then you are actually going to precipitate or isolate the genomic DNA by the alcohol precipitation. So, you are going to in the third step, you are going to precipitate the genomic DNA by the absolute alcohol. And after that, you are actually going to get the DNA and as well as protein and as well as the lipids. So, then you are actually going to do the purification step. So, you are going to extract the things with the help of the chloroform and phenol chloroform isomer solutions.

And when you do that, you are actually going to get the two phases. You are going to get the aqueous phase and you are actually going to get the organic phase. In the organic phase, you are actually going to have the proteins whereas the lipids, whereas in the aqueous phase, you are actually going to have the DNA or the genomic DNA actually. And then you can collect this and again, you are going to precipitate the DNA with the help of the absolute alcohol and that is how you are actually going to get the pure mammalian DNA. If you analyze this genomic DNA onto agarose gel, we are not discussing about the agarose gel in this particular course.

What you will see here is that it is actually going to run as an intact band and it will run very close to well actually. This is a well where you have loaded actually. Why it is so because the genomic DNA is very big and it is actually quite. So, it is actually going to be slow, run very slowly. So, genomic DNA is actually going to be analyzed on 0.

28% agarose and a good preparation of genomic DNA gives an intact band with no visible spheres. Now, once you isolated the genomic DNA, you isolated the genomic DNA, you are actually should do the estimation. You should know what is the amount of DNA what you have isolated. So, if you want to isolate the genomic DNA, you have two choices.

One, you can actually have the absorbance at 280 nanometer or the other is you can actually be able to do the colorimetric method.

So, you can actually do the absorbance at 280 nanometer. So, what you can do is you can take the small amount of DNA and then you actually add the buffer. So, you can actually be able to. So, what you do is and you know that the DNA is absorbing very strongly at 260 nanometer and RNA also. So, RNA and DNA both absorbs very strongly at 260 nanometer. So, what you can do is you can take the buffer and first take the absorbance at 260 nanometer and that is going to be the control reading or I will say blank.

So, it is actually going to be considered as zero reading and then what you do is you add the buffer and you add the small amount of DNA. For example, add the 2 microliter of DNA and again, I will take the absorbance. So, it is actually going to show me the absorbance of for example, 0.

15. So, this is absorbance what I got at 260 nanometer and I can thus convert this and get the concentration of DNA. So, that you can do in a spectrophotometer or you can actually be able to use the nanotopes. So, we have prepared a small demo clip and where the students are actually going to show you a small demo clip. So, that you can be able to determine the DNA concentration with the help of the absorbance at 280 nanometer. Today, we are going to estimate DNA concentration using UV visible spectroscopy. One of the most common methods for DNA concentration detection is the measurement of solution absorbance at 260 nanometer due to the fact that nucleic acids have an absorption maximum at this wavelength.

For this experiment, we need DNA for standard curve preparation, distilled water, DNA sample of unknown concentration, micro pipettes, tips, cuvette and spectrophotometer. According to this table, we will prepare different concentrations of DNA solution for the standard curve. After preparing different concentration of DNA solution, we will measure the absorbance at 260 nanometer using spectrophotometer.

We will take the absorbance of the blank. Now, we will take the absorbance for the 5 microgram per ml concentration. These are the absorbance values. From the absorbance value, we have plotted the graph and we have got the regression equation. Our absorbance for unknown sample was 0.478 and this value is the value of the Y.

Substituting the Y value in the regression equation and solving it will give the X value which is our unknown concentration that is 22.319 microgram per ml. Now, the second method is you can actually be able to do the DNA estimation with the help of the chloromyrtic method and that method is called as estimation of DNA by the diflamine

reactions. So, diflamine is a chloromycytic reagent when it reacts with DNA. So, if it reacts with diflamine, it is actually going to give you the blue colored complex and that blue colored complex you can actually be able to give you the absorbance and that is actually can be used for determining the total DNA content what is present.

Now, the question comes why we are doing this instead of the absorbance at 260 nanometer. The answer to this question is that the absorbance at 260 nanometer is a quick method and it actually gives you the quite reliable results, but it is not very, very quantitative. It will not go to give you the absolute correct answer and that is why there is a chloromycytic method in case you want to verify because if you are doing at 260 nanometer absorbance, there are other molecules which also can contribute into the reactions. So, what is the principle of the diflamine reactions? So, the deoxyribose, the sugar part in DNA is in the presence of acids forms the beta hydroxy linoleumaldehyde, which reacts with the diflamine to give a blue colored complex which is shot of absorbance at 595 nanometer. In DNA only the deoxyribose of the purine nucleotide reacts.

So, the value what you are going to obtain represent the half of the total deoxyribose. So, what you have is you have a DNA, DNA has pyrimidines and the purines also. So, when you are going to put them into the acidic reactions, it is actually going to have the purines and then you also going to have the sugar part. This sugar in the presence of sulfuric acid, it is actually going to react with the diflamine to form the beta hydroxy linoleumaldehyde. And then the beta hydroxy linoleumaldehyde is actually going to react with the diflamine and the diflamine is actually going to give you the blue colored complex and this blue colored complex is going to absorb very strongly at 595 nanometer.

What are the metal logic? The material requires, you actually requires the spectrophotometer and a water bath. So, you require a boiling water bath. Remember that you actually require a water bath which actually can maintain 100 degree Celsius. Then you require the chemicals, you require the standard DNA solutions, diflamine reagents, DNA sample. You require the citrate buffers, you require the acetic acid, concentrated sulfuric acid and ethanol.

The glassware, you require a test tube, pipettes and graduated cylinder. Then the procedures, so you are actually going to prepare the reagents. So, you are going to prepare the diflamine reagents and while you are preparing the diflamine reagents, remember that these are the reagents where you are actually going to have the glycolytic acid and you are also going to have the concentrated sulfuric acid. So, you should be taking very careful, very lot of cares and the reagent has to be stored in a dark glass bottle. So, on the day of use, prepare a fresh solution of methanol and you are going to add the things into like 1 ml of methanol in the 50 ml of water and you add the 0.

5 ml of the solution to each 100 ml of the diflamine reactions. You have to always be very cautious because you are actually going to deal with the concentrated sulfuric acid and concentrated glycolytic acid. So, you always use the eye wear protection and you use the fume cupboard for these reagents. And diflamine reagent is also very harmful.

So, if ingested or inhaled, may irritate the skin or eyes and it comes out into the contact with them. Now, we are going to set up the assay. So, you are going to prepare a series of dilution of a standard DNA starting from the 0.25 mg per ml stock in a saline-seated buffer to give a concentration of 50 to 500 microgram per ml. You prepare all the sample in duplicates, then to 2 ml of each of these blank, standard and unknown, you add the 4 ml of diflamine reagent and mix. Tube 1 is used as a blank and tube 2, 2, 7 are used as a concentration of a standard calibration curve, whereas tube 8 to 11 are for the unknown samples.

This anyway you are going to see in the table, right. Then you incubate all the tubes into a boiling water bath for 10 minutes, cool the temperature and read the absorbance at 500 nanometer. Then you can actually be able to make a calibration curve of the absorbance at 500 nanometer versus the consideration of the DNA. And this is a table, this is the recipe table what you are going to use.

So, the first reaction is actually the blank reaction where you have not added the DNA. So, this is the minus DNA reactions. So, this is actually going to be the blank. So, whatever the reaction you are going to, whatever the values you are going to get that has to be subtracted from this value, right. So, this is the value what you are going to get, right. This is the average of this and this value has to be subtracted and when you do that, you are actually going to get, this is the corrected absorbance value at 525 nanometer.

And these are the standard reactions what you are going to use, right. And these are the unknown samples. So, DNA what is present in the unknown samples. And then using these reactions, using these values, you can be able to make a calibration curve. And using this calibration curve, you can be able to determine the DNA concentrations into the unknown samples. So, this is the standard curve what you are going to prepare where you are going to have the corrected absorbance value onto the y-axis, right. This is the y-axis and the concentration of the DNA or amount of DNA onto the x-axis. And then you are going to get the calibration curve and you can actually be having two options, either you use the equations and you can be able to determine the concentration of the unknown samples, right.

Or you can actually be able to use this calibration curve to determine the unknown

samples. So, to show you all these, we have prepared a small demo and with this small demo, you can be able to understand how to prepare these reactions and what are the different places where you are supposed to take the precautions. Because you are dealing with the corrosive samples. So, you should be very, very careful.

Hello everyone. In this video, we are going to demonstrate how to estimate the concentration of DNA using the diphenyl amine method. The basic principle in this method is that the deoxyribose in DNA or the purine nucleotide in the presence of sulfuric acid going to form beta-hydroxy-levinol-dehyde. That in turn going to react with the diphenyl amine forming a blue colored complex with absorbance at finite different nanometer. So, here the material which will be requiring are the standard DNA solution prepared in citrate buffer of 250 microgram per ml. The diphenyl amine reagent, the saline citrate buffer, the test tubes for the standard curve preparation of 6 standards, 3 unknowns as well as the buffer blend.

And also we have been needing the hot water work as well as the spectrophotometer for the absorbance. So, coming to the procedure to prepare the standard curve, we need to add the known concentration of DNA in each of the standard tube. So, in standard 1, we will be adding 50 microgram of DNA, in standard 2 100 microgram of DNA, in standard 3 200 microgram of DNA, in standard 4 200 microgram of DNA, in standard 5 400 microgram of DNA and in standard 6 500 microgram of DNA. So, to add this particular concentration of DNA in each of the test tube we need to we have already have the standard DNA solution of 250 microgram per ml. So, for 50 microgram of DNA to be added in standard one we need to add 200 microliter of the standard DNA solution with 1800 microliter of water into the standard one test tube. So, likewise for 100 and 200 and for 100 200 200 like that we will be adding the DNA non DNA concentration along with the distilled water to make up the volume 2 by ml in each of the standard test tube.

And now we will be adding the known concentration of DNA into each of the standard test tube. So, for the standard one we will be adding 200 microliter of standard DNA solution of 250 microgram per ml to make it into a known concentration of 50 microgram. And we will be adding the distilled water to make up the concentration of 2 ml. Likewise for other standard tubes with the known concentration of DNA we will be making up to 2 ml in each of the standard tubes. Now, after preparing the known concentration of the DNA in each of the standards along with the unknown now we will be adding the 4 ml of DPA reagent to each of the test tube including the gland.

Likewise we are going to add to all of the standards as well as the unknowns. So, we will be adding the DPA reagent to each of the test tubes to make up a total volume of 6 ml in each of the test tube. Now, we have added the DPA reagent of 4 ml each in all of the test

tubes. Now, after adding we are going to incubate all of the test tubes in hot water bath that we are going to put it for 10 minutes. After incubating the samples for 10 minutes at 100 degree centigrade now we need to let the samples to cool down to room temperature.

Now, we could see the blue coloured complex formed in the standards as well as in the unknowns. So, taking from the standard 1 to 6 we could see as the concentration was increasing the intensity of the blue colour was also increasing. Now, to record the absorbance we need to check at finite defined nanometre in the spectrophotometer. Now, we are going to measure the absorbance at finite defined nanometre using the spectrophotometer. So, this is the spectrophotometer device where we are having two cuette holders one is for the blank other is for the test samples.

So, first we need to set the absorbance at finite defined nanometre and I am going to take the blank in one of the cuette and place it in the cooler. So, we are going to take the standard one in another cuette and place it in the another cuette holder. This would be constant and for the standards we would be changing from second third to third is the unknown samples. Now, it is been measuring. Here are the absorbance values taken twice for each of the sample including the unknowns to reduce the error percentage by taking the average of two values.

The last column gives us the corrected OD after nullifying the blank from each of the cuette and unknown samples. Now, by plotting the standard curve with absorbance on y axis and quantity of DNA on x axis we have determined the unknown concentration of DNA using the equation of slope. Here the obtained concentration is usually half since the purine nucleotides only forms the blue coloured complex after reaction with the diphenyl amine reagent in the presence of strong acid.

So, doubling the obtained concentration for each unknown gives us the actual concentration of DNA. Like for unknown one it is around 0.4 microgram. So, by doubling it you get around 0.8 microgram. So, this is all about the DNA. We have some more aspects what is related to the deoxyribonucleic acid which we are going to discuss in our subsequent lectures. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 03
Basics of Biomolecules
Lecture-12 Biomolecules (Part 2: DNA Sequencing)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering, IIT Guwahati. you are purifying the human genome or if you are purifying the human genome, it should not be the case that you actually isolated something else, right. And that information only you will get when you are actually going to do the sequencing of this particular DNA, right. And that is very, very relevant when you are actually working with us, you know, the recombinant clones, right. For example, you generated a recombinant clones, then it is important that you verify that DNA with the help of the sequencing reactions.

So first thing is how you can be able to determine the purity of the DNA, right. So the purity of the DNA can be determined by spectrophotometer. So you can you know that the DNA and RNA absorbs at 260 nanometer, whereas the protein absorbed at 280 nanometer. But if you see the curve, right.

So if you see the curve, if you what you will see here is that it 260 the DNA and RNA are going to absorb, but they also absorb at 280 nanometer, right. So this is for DNA and RNA. Whereas for the proteins, it is actually going to be like this. So this is for the protein, right. Now what you see here is that this is showing a lambda max for the nucleic acid, right, for the nucleic acid, whereas this is the lambda max for the protein.

But it does not mean that the protein is not contaminating or not, you know, having any absorbent at 260 nanometer. And that is why it is important that we should calculate a ratio of 260 to 280 ratio. So what you can do is you take the absorbance not only at 260 nanometer, but at 280 nanometer. And that is actually going to tell you the purity of your sample. For example, if you calculate the 260 to 280 ratio, and if it comes at 1.

8, right, then it is a pure DNA, right. If 260 to 280 ratio is less than 1.8, then there will be a contamination of the protein, which means this component is now increasing and this component remains the same, right. So, this means it is actually going to, the ratio will actually going to lower down, right. This means, see, what will be the, if it is a 50-50 contribution, right, then the ratio is going to be 1, right.

Because then the absorbent at 260 and 280 are actually going to be equal, right. So that

time the contamination would be 50%. But if it is less than 1.8, then there will be a contamination of protein. Now if 260 to 80 ratio is more than 2, right, which means you are actually going to have very high absorbance at 260 nanometer, that means there will be a contamination of RNA into the DNA pair.

Now this is very important to understand. You know that the DNA is double standard, which means you have two strands and the bases are inside. So, whereas the RNA is mostly been single standard, right, which means this is the RNA with the bases. Now, if you see the very carefully, the bases within the DNA are being protected within the DNA structure, right. And because they are not exposed to the outer environment, because one base is, you know, put next to each other, they are actually going to show the lesser quantum yield and lesser excitation to the light, okay.

And because of that, they are actually going to show you the lower absorbance compared to the RNA molecule because RNA the bases are exposed to the water and exposed to the outer environment. And because of that, they will actually going to make the absorbance more, right. And because of that, the RNA is actually going to show you the more absorbance compared to the DNA. So, even if the DNA is pure and if there is a RNA contamination, the RNA is actually going to show you the 260 reading more. So, if the 260 reading is more and 280 is same or equal, then the ratio will actually above 2, right and that is how it is actually going to give you the indication that there is a contamination of RNA species.

Now, this is all about the DNA purity, right. Now, the second point is about the sequencing. So, DNA sequencing, right. So, DNA sequencing historically, there are two methods of DNA sequencing with a similar principle of breaking the DNA into the small fragment followed by the separation and analyse them on a high resolution electrophoresis gel. So, if you want to sequence any biomolecule, you are supposed to do this, right.

For example, if this is the DNA, what we want to sequence, what we can do is we can just split this into small, small, small, small fraction, right. And then we can actually be able to sequence. So, it is actually the same rule as divide, right and sequence. Because it is easy to then manage these small fragments, you cannot actually manage a 3 KB DNA, but you can easily manage 100 base pair DNA because then it is easy to manage, right. So, what you can do is you amplify this particular sequence with a modified base.

So, when you amplify this with a modified base, wherever you are actually going to have the modified base, it is actually going to break, right. And that is how it is actually going to give you the small fragment. And you know where it is actually going to break.

For example, if I take the modified base for A, it is going to break here. It is going to break.

Yeah. And if I take the modified base like G, then it is going to break here. It is going to break here and something like that. And then I can analyze these sizes of the DNA on a high resolution electrophoresis gel and that is how it is actually going to give me the complete sequence. And if you follow this method where you are going to use the modified base, then this sequencing method is called as Sanger sequencing method. And for this sequencing method, the Sanger got the Nobel Prize.

Now other option is that you use the chemical reagents, right, you use the chemical reagents which are actually going to attack the bases. So, you can have the reactions for A, you can have the reaction for T, you can have the reaction for G, you can have the reaction for C. So, what will happen is it is going to break after A, T, G, C like that. And that is how you can actually be able to separate these fragments and that is how you are actually going to get the information about the sequencing. And if you use that method, then it is actually going to be called as Maxim-Glibert method.

So, let us first discuss the Sanger sequencing method and then we are actually going to discuss about the Maxim-Glibert method. So the Dideoxy chain termination method or the Sanger's method. This method is originally been developed by the Frederick Sanger in the year of 1977. In this method, a single standard DNA is used as a template to synthesize the complementary copy with the help of a polymerase in the presence of T nucleotides. The polymerase reaction contains a primer and the nucleotides, three normal nucleotides and a 2 prime, 3 prime dideoxy nucleotide triphosphate which is a modified nucleotides.

When the DNA polymerase utilizes the DDNTPs as nucleotides, it gets incorporated into the growing chain, but chain elongation stop at the dideoxy as the dideoxy due to the absence of 3 prime hydroxyl group. In the typical sequencing reaction, you are going to run the four different DDNPs are taken into the four separate reaction and analyze on the higher resolution polyacrylamide gels, the ratio of NTPs and the NTPs is adjusted so that the chain termination occurs at each position of the bases in the template. So when you do the dideoxy chain termination method, you can actually have this right. So this for example, this is the region which you want to sequence right. So you're actually going to have the primer.

So in the step one, a primer is added and annealed to the 3 prime of the DNA helix right in a template. In the step two, the radio labeled ATP is used to label the primer. So you are actually going to label the primer so that you know what will be the fragment. So you

can actually be able to identify this fragment onto the autoradiograms. Then the step three, the polymerase reaction is divided into the four reactions.

So you can have the four reactions, you can have the A reaction, you can have G reaction, you can have C reaction and you can have D reactions. So in the A reaction what you have, you have the A DD ATP actually, so dideoxy ATP. In the G reaction, you're going to have DD GTP. In the C reactions, you're going to have DD CTP right. And in the T reactions, you're going to have DD TTP right.

This is what it is showing here. So you can have the A reactions, you can have the D reactions, you can have G reactions, you can have C reactions and all the four reactions, you are actually going to be load onto the sequencing gel and then you are actually going to analyze them with the help of the rate autoradiogram. So in the step four, the DNA synthesis continue until the terminated by the incorporation of the specific dideoxy nucleotides right. Because the dideoxy nucleotides does not contain three prime hydroxyl groups. So it will actually going to terminate the chain elongations.

Then you are going to a chase of the polymerase reaction is performed in the presence of high concentration of NTPs to obtain the all non terminated sequence into the high molecular weight DNA. This high molecular weight sequence will not enter into the sequencing gel. So because the pore size what you are going to adjust in of the sequencing gel in such a way that these high molecular weight DNA is not going to enter because this high molecular weight DNA will not going to provide you any information about determinations and they are actually going to make the analysis more complex. So after this, you are going to have the four reactions, you can have the A reactions, you can have D reactions, you can have G reactions, then you can have C reactions. So the way it goes that you are actually going to have the A reactions right, then you can actually have the T reactions.

So from the A you are actually going to have the T right and from the T you are going to have the T. So you are actually going to read from the bottom okay. So for example, you are going to have A, T, T, A, G, then you are going to have A right, then you are going to have C, then you are going to like that. So if you go like this right, you go like this, then you are going like this, you go like this. So you have to read in the reverse direction.

So the smallest one you are going to put first. Then last you are going to do like this, third one like this, fourth one, fifth one, sixth one, seventh one, eighth one like that. Like you have to go from the bottom and you have to keep reading and keep putting the sequences like this and ultimately you are going to get the sequence of the DNA. Now

let us talk about the Max and Gilbert method. So the Max and Gilbert method actually relies on the chemistry part right.

So it is actually going to utilize the different types of reagents which are actually going to be basis specific. So you can have the A reaction, G reaction, C reaction, T reaction and so on. And that is how it is actually going to do the same thing what we are doing, what the Sanger has done with the help of the enzyme. But here you are actually using the different types of chemical reagents. So this method was discovered by the Max and Gilbert in the year of 1977 which is based on the chemical modification and subsequent cleavage.

In this method a 3 prime or 5 prime radionuclide DNA is treated with a basis specific chemical which is randomly cleaved the DNA at their specific target nucleotides. These fragments are analyzed on a high-resolution polyclomer gel and the autoradiogram is developed. The fragment with the terminal redolibut appear as a band in the gel. So the chemical reaction what are going to be performed in 2 steps. First you are going to have the basis specific reaction and the second step you are going to have the cleavage reaction.

So the basis specific reactions, first you are going to have the basis specific reaction. So different basis specific reactions are used to modify the target nucleotide. So reaction 1 you are going to have the dimethyl sulfate or DMS which is actually going to modify the N7 of the guanine and then open the ring between C8 and C9. This is going to be called as G reactions. Then in the reaction 2 you are going to use the formic acid and act on the purine nucleotides.

So it is actually going to act on G and A by attacking on to the glycosidic bond. Then you have a reaction 3 which is going where you are going to have the hydrazine and which is actually going to break the rings of the pyrimidines. So it is actually going to be not specific for a particular base but it is only going to be specific for pyrimidines. So it is going to be called as T plus C reactions. Then you are going to have reaction 4 where in the presence of salt it breaks the ring of the cytosine.

So it is going to called as C reaction. So basically you are going to have 4 reactions. One is called G reaction, other one is called as A plus G reactions. You are going to have C plus T reactions and you are going to have C reactions. So you are going to take the radiolabel DNA, you are going to add these reagents and that is how you are going to have the G reactions, A plus G reactions, C plus T reaction and C reactions. And then you are going to have the cleavage reactions.

So after the base reactions pyrimidine is added which will replace the modified bases and catalyze the cleavage of phosphodiester bond next to the modified nickel. So once you have done the reaction after the G, after the G, after G and C, after C it is going to be cleaved and that is how you are going to have the fragments. So the fragment, so you are going to have the G reactions, G plus A reactions, C plus T reaction and C reactions. So imagine that if this is the DNA sequence which you want to sequence.

And here also exactly the same way. You have to go in the reverse direction, but remember that when you have the bond band between G and G plus A then it is actually going to read as G. So it is for example here you have the 2 bands. One is in the G reaction, the other one is in the G plus A reaction. So it is not going to be read as A, it is actually going to be read as G.

So from here this is going to be read as G. Then you have this, so you actually will go in the reverse direction. So the fragment in J line is read as G whereas the fragment in the G plus A but absent in G is read as A. Similarly fragment in C is read as C whereas the fragment present in T plus C but absent in C is read as T. Same is true for this C. If the 2 fragments are present in same distance, of same size then it is actually going to read as C rather than T actually.

But if the fragment is absent in C but it is present in T then it is actually going to read as T. For example this one. So this one the fragment is absent in G but it is present in G plus A. So this is actually going to be read as A.

Same is true here. For example this is the T plus C. So this is actually going to be read as T and that is how if you go from the reverse direction you are actually going to be reduce the sequence at the end and that is what it is actually going to be the sequence of that particular DNA. So the way you are going to read this sequence is that you are actually going to read the lowest with the band and then you are actually going to go to the higher bands. So this is all about the assessment of the DNA quality. So what we have discussed, we have discussed about the purity of the DNA and we have also discussed about the DNA sequencing. While we were discussing about the purity of DNA we took the help of the spectrophotometer and when we were discussing about the DNA sequencing we discussed about the Sanger's method and we have also discussed about the Max and Gilbert method.

So with this I would like to conclude my lecture here. Our subsequent lecture we are going to discuss some more aspects related to the biomolecules which are relevant for the molecular biology. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 03
Basics of Biomolecules
Lecture-13 Biomolecules (Part 3: RNA)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati and what we were discussing, we were discussing about the different types of biomolecules in the course molecular biology. Both of these biomolecules have very crucial roles in running the different types of metabolisms or they are contributing in one or other way into the different types of molecular biology processes. So, what we have discussed, we have discussed about the DNA. In the previous lecture, we have discussed about the DNA and DNA is the crucial biomolecule which is required for the genetic information. So, it carries the information from the one generation to another generations and then we have also discussed that the DNA which is going to be involved into a process called replications and the transcriptions.

So, with the help of the transcription, the DNA is actually going to synthesize the RNA and this RNA is then actually going to participate into the reaction of the protein synthesis in a process which is known as translations. So, the enzyme which is actually going to synthesize the RNA from the DNA is known as the RNA polymerase. So, RNA polymerase is actually going to read the DNA and it is actually going to synthesize the RNA and this process is called as transcription and then RNA is actually going to participate into a series of events which is known as translation to synthesize the protein. And so, in today's lecture, we are going to discuss the properties of the RNA and the structure of the RNA and how you can be able to isolate the RNA from the cell and how you can be able to characterize the RNA so that and estimate the RNA and so on.

So, why it is important for us to understand the RNA because the RNA is mainly been responsible for synthesis of the protein. So, if you want to do an experiment related to the expression studies and other kinds of studies, then you are supposed to study the RNA. Now, when we talk about the RNA, we are actually going to talk about the three different types of RNA. So, we have three different types of RNA, we have the transfer RNA or commonly known as tRNA, so this is called as transfer RNA. Then we also have the ribosomal RNA or it is called rRNA and we also have the messenger RNA or it is called as mRNA.

The mRNA is actually going to provide the message or actually it is going to provide the informations in which sequence we are actually going to add the amino acids, so it

actually going to provide the information of the synthesis which means which amino acid I should add like that kind of information. So, for example, if I want to start writing a letter, I have to first know what is the sentence. So, if I know the sentences, then my brain is actually going to read that sentence and that sentence is nothing but this messenger RNA. And then I am going to bring A, B, C, D like that, so that A, B, C, D information is this. Now you are actually going to read the help of the transfer RNA and the ribosomal RNA.

So, transfer RNA is actually going to bring the amino acids in the same sequence what the sequence is given here, right. So, if it says you bring the alanine, right, it is actually going to bring the alanine. If it is going to say the methionine, then it is actually going to bring the methionine. Then who will bring the, so this is actually going to bring the amino acids. Then these amino acids are actually going to be joined by the ribosomal RNA, right.

And you know their amino acids are actually going to join by a bond which is called as peptide bond. So, basically the job of the ribosomal RNA is to form the peptide bond between the A and B, right. And that is how it is actually going to start synthesizing the protein molecules, right. So, these are the some basic or the brief overview of that function of these RNA species. One it is actually going to provide the message.

So, it is actually going to provide the information of synthesis, right. In what sequence I should add the amino acids and also going to provide which amino acids. And then the transfer RNA is actually going to bring that particular RNA and the ribosomal RNA is actually going to collect the information from the transfer RNA and messenger RNA and that is how it is actually going to join the amino acids by the help of a peptide bond. And that is how you are actually going to have the sequence and it is actually going to have the information, okay. So, this is just alphabets actually.

There is no amino acid with the B, right and so on. So, today in today's lecture, we are not going to cover the structure of transfer RNA or the ribosomal RNA. In today's lecture, we will only focus on to the messenger RNA. We will take up the structure of transfer RNA and ribosomal RNA when we are actually going to discuss about the translation because all of these three messenger, all of these three RNA species are actually going to participate actively into the translation process. So, we are only going to focus today, in today's lecture, we are only going to focus on to the messenger RNA.

In our subsequent lecture, we will focus on to the transfer RNA and ribosomal RNA when we are going to discuss about the translation. So, we will talk about the structure of the messenger RNA. So, messenger RNA is actually as I said, it should have the

information of the synthesis, right. So, it is going to provide the information of the synthesis of protein, which it actually going to take up from the DNA, right. So, that information is originally being present in DNA, but that information will then going to be taken up by the messenger RNA.

And so, messenger RNA is having the three distinct part. It is actually going to have the, so this is the 5 prime end and this is actually the 3 prime end, okay. And remember that we have discussed in detail about the RNA structures when we or the composition of the RNA when we are talking about the DNA, right. So, RNA is actually also a poly nucleotide molecules, RNA is single stranded in majority of the cases and RNA is also going to have the phosphodiester linkages and phosphate and black bone. So, RNA is also going to be made up of the sugar, phosphate and base, right.

And as far as the base is concerned the RNA is actually going to have the A, G and U and C, it does not have the T, right. So, there is no T present, okay. So, T is absent in the case of RNA, whereas the T is present in the DNA. Instead of T, it is actually going to have the uracil, okay. And that is the basic difference between the RNA and the DNA.

Other than that, it is going to be single stranded. So, there will be extensive secondary structures what are going to be present in the RNA species. And as far as the structure is concerned the RNA is going to have the 5 prime cap. So, this is the cap which is actually going to be protect the RNA sequences because RNAs are very, very susceptible for the RNA molecules. And then it is actually going to have the 5 prime UTR.

So, 5 prime UTR is a place which is actually going to provide the docking site for the RNA polymerase. In detail actually how the RNA polymerase is going to sit and how it is actually going to recognize the promoter regions and all that, that actually we are going to discuss when we are going to talk about the translation. So, in the 5 prime UTR you are going to have the promoter, right. And there is a definite composition of the promoter. So, it is going to have the data box, it is going to have the minus 10 region, minus 35 regions and so on.

So, all that I think can be discussed when we were discussing about the transcription and translation. So, promoter can be of strong promoter or it can be of weak promoters. So, promoters are actually going to provide a docking site for the translations initiation site. So, it is going to provide the translation initiation site and it is going to allow the sitting of the RNA polymerase. And so, it is going to provide the docking site for the protein synthesis.

So, it is going to provide a docking site for the ribosomal machinery. And this was

going to be a promoter. So, it is actually going to be a strong promoter or the weak promoter and it is actually going to provide the docking site for the translational machinery. Next to this, you are actually going to have the coding sequence. So, this is a region which is going to be a coding sequence.

So, this is the region which is going to give you the protein. So, it is actually going to provide the information in terms of the genetic code and these codes are actually going to be read by the ribosomal machinery and as well as by the tRNA and that is how it is actually going to have the help you in the synthesis of the protein. So, genetic information is encoded in the face of genetic code and each genetic code is actually going to be corresponding to the amino acids or these amino acids are then going to be added into the ribosomal machinery by the help of the peptide bond and that is how it is actually going to synthesize. And then you also have the 3 prime UTRs. So, 3 prime UTRs are actually going to be the regulatory side which actually going to provide the regulation of this whole translation process and then at the 3 prime end you are actually going to have the polyadenylation side.

So, polyadenylation is very important because polyadenylation you can actually have the addition of the ACE starting from the 50 to 200 and depending upon the polyadenylation you are actually going to decide the age of the messenger RNA because this is actually going to chewed because remember that from this side by a RNase. So, if the RNA is chewing this amino acid this messenger RNA because messenger RNA is going to be present in the cytosol inside of the nucleus. So, RNA is actually going to be synthesized by a process of transcription inside the nucleus then it is actually going to be transported outside into the cytosol and then it is actually going to provide a docking site for the protein synthesis machinery which means it is actually going to allow the assembly of the ribosomal machinery like the small subunit and the large subunit and that will happen on to the 5 prime UTR and then it is actually going to synthesize the proteins. But how long this amino acid this messenger RNA is going to remain active into the cytosol that will be decided by the 3 prime poly A tails. So, this region is called as the poly A tail and this poly A tail is going to be or the ribosomal RNase which are very active within the cytosol are actually going to chew these RNA from the 5 prime, 3 prime end and the moment they are actually going to be chewing up chewing up like this they once they hit the coding sequence then they are actually going to start you know the this messenger RNA will not be useful for the synthesis or for the synthesis of the protein because now it is actually going to start synthesizing the you know the cryptid protein or the truncated proteins and that may not be good for the cell.

So, this length of this polyadenylation or length of the poly A tail will actually going to inside the edge of this particular messenger RNA or I will say the stability of this

messenger RNA within the cytosol. So, there are messenger RNA where you have the very huge number of A's and they will actually go to remain in the cytosol for very very long time so that they will be keep expressing the protein. So, some of these messenger RNase are belonging to the housekeeping genes. For example, you have the messenger RNA for actin, myosin, LDH. So, these are the proteins which are required in a you know in a there is a huge demand of these proteins and that is why they are supposed to synthesize and since they are housekeeping genes or they are actually housekeeping proteins their level will actually go to determine the health of that particular cell and that is why they will actually go to have the huge number of poly A tail or huge number of amino acids A residues into the poly A tail.

The poly A tail is a very very interesting tool because it also provide the stability to the messenger RNA. At the other hand, it also can be a tool to purify the messenger RNA from the cytosol. So, that we can do with the help of a affinity column and that is what we are going to discuss now. So, we are going to purify the messenger RNA from the cytosol. So, there are two methods by which we can be able to isolate the messenger RNA.

One we can actually use the affinity column and we can use the affinity DT columns and the other approach is we can actually be able to use the trizone method. So, in an affinity DT column what we have is we have a linker actually and it is attached to the beads and this linker has the T residues which means it is actually has the thymine. So, when you have a thymine you know that A is always making a pair with T and G is always making a pair with C. So, utilizing this information if you have the A's what is present onto the messenger RNA. So, what you can do is you can actually be able to have the beads and you can actually be able to put the beads into the cytosol.

So, what will happen is that you are actually going to have the binding of A's which are be a part of the polyadenylation tail and it is actually going to have the messenger RNA. So, this is the messenger RNA and it is going to have this right. And ultimately what you can do is you can actually be able to do the elution and at the end what you are going to have you are going to have this eluted right. So, this is actually going to give you the complete pool of messenger RNA. This is another method which is called as triazole method where you are actually going to use the triazole and that is how you are actually going to isolate the messenger RNA.

So, both of these methods we are going to discuss how you can be able to isolate the messenger RNA from the cytosol. So, for first method, so this is the method 1 right and this is the method 2. So, for the first method what you are going to do is you are going to test grow the cells right. So, these are the target cell.

You can have also the tissue right. So, depending on what kind of material you are actually. So, if it is a tissue then you are actually going to grind the tissue. So, that it is actually going to give you the single cell suspension right. Sometime you might have to use the enzymes and other kinds of treatments.

So, that we are not discussing there. So, but if it is starting with a tissue for example, you started with liver right. So, if it is started with the liver then it has to be grind fine with the cell mortar or the homogenizers and then the liver is actually going to give you the single cell suspension and then from the single suspension you are actually going to use the same way as you are actually going to use the cell from the cell culture. So, you are going to put them into a lysis buffer. Mostly the lysis buffer contains the SDS and also contains the protein SK okay. And it also contains the sometime SDS or sometime TITON-X100 okay.

So, it is basically going to contain the detergent and the protein SK and it is also going to have the binding buffers. So, you lysed the cells under the lysis and the binding buffer and then you are going to, so this is what we have shown here right. If you have a tissue you can just do the homogenization so that you are going to have the single cell suspension and then you can incubate this. Once you incubate this it is actually going to lyse the cells and you are going to have the cell lysate okay. So, with the cell in the cell lysate you can actually you can do the spinning at for example 1000 rpm.

So, that we are actually going to remove the nucleus and then it is actually going to give you the cell lysate because the nucleus is useless because it actually going to increase the contamination. So, nucleus if you remove the nucleus you are actually going to get rid of the DNA right. And then you take the messenger RNA you put it into the binding buffer and then you are actually going to have the oligo-dt beats and as I explained the oligo-dt beats are that you are going to have the aero beats and then it is also going to have the linker that linker is actually going to have the T residues attached to it okay. This means it is these linker are actually going to have very strong and specific binding for the a residues okay. So, when you do that the messenger RNA what is present in this particular cytosol.

So, it is not specific for a particular messenger RNA it is actually be responsible for all the messenger RNA and that is how it is actually going to bind the messenger RNA. So, this is the messenger RNA. So, they will interact with each other and then you are going to do a washing with the buffer because there could be some non-effective interaction. So, you can do a washing with the buffer that washing can be done with the you know the buffer with salt right.

So, you can actually add some salt. So, that you are actually going to reduce the non-specific interactions and then you are actually going to do the elusions okay. So, once you have the pure sample you are actually going to do the elution. So, elution can be done with the so you at this step you collect the beats and then you are going to discard the thropanactin and then you are going to do the elution. So, you can add the for example you can add the poly T or you can actually add the thiamine right and then you can suspend that into the elution buffer and the elution buffer is going to allow or it is actually going to break the hydrogen bonding between the poly T tail which is attached to the beats versus the poly A tail which is present onto the messenger RNA and that is how the messenger RNA is going to be eluted and then you can actually take this pure messenger RNA for the further downstream applications like RT-PCR and you can use that for other kinds of applications. So, this is exactly what people were doing when you are actually asking them to do the COVID testing okay.

So, they were taking your saliva and other kinds of samples and then they were doing this process to isolate the messenger RNA and then they were doing the RT-PCR with the help of the primers for COVID and that is how they were saying that if it is a they were getting the amplifications of the DNA for the of the cDNA then they are actually saying that it is COVID positive. Anyway, that is separate part that anyway we are going to discuss when we are going to talk about the real-time PCR and reverse transcriptase RT-PCR and we will also going to take up how you can be able to use the RT-PCR for these kind of applications. So, this is the first method where you are actually going to use the affinity column to purify the messenger RNA from the cell lysate or the tissue. Now let us move on to the next method and the next method is called as the Trizole method. So, the Trizole method RNA isolation by the Trizole method, this RNA isolation by the Trizole method uses the Trizole which is also called as the tri reagent for the isolation of the total RNA.

Trizole is a mixture of guanidine thiocyanate and phenol which effectively dissolve the DNA, RNA and the protein on homogenization or the lysis of the tissue samples. After adding the chloroform and centrifugation the mixture separates into the three phases with the upper clear aqua phase containing the RNA, interphase containing the cell debris and the lower is the organic phase having the protein and the lipids. The next step in the extractions are the washes and the precipitation of the RNA. The first part of the protocol from the homogenized tissue in Trizole to the point of an RNA palette in 75% ethanol take less than an hour. The RNA can then be stored for long period of time at minus 20 degree Celsius.

So RNA is very stable when you are isolating with the Trizole method and putting it

into the 75% ethanol. The same protocol can be used for RNA extraction from the cell culture. So if you want to remove the DNA you can actually be able to treat the sample with DNAs and that is how it is actually going to remove the DNA part. So this is what it is actually going to say that if you have the grinded adipose tissues, for example, this is the tissue. So in the step one, you are actually going to add the reagent and that is how you are going to vertex then you are going to wait five minutes at room temperature and then it is actually going to give you the aqua phases and you are going to have the different types of buffers and what you are going to see here is that when you are going to have the phases, you are going to have two phases.

One is the RNA phase, the other one is going to be chloroform phase and in this phase you are actually going to have the protein plus lipid. Whereas in the aqua phase, you are going to have the messenger RNA and that you can actually be able to. If you transfer this aqua phase, then you can actually be able to use that by precipitation with the 75% ethanol and you can actually be able to air dry this pellet and then you dissolve this into a RNA 3 buffer and that is how you are going to have the RNA. So let us see what are the different methods or different protocols, right. So this procedure is very effective for isolating the RNA molecules of all types from 0.

1 to 15 kB in length. However, there are commercial kits that enable the simple RNA extractions using a column that binds the RNA and so on. That is anyway we have discussed, right. So what are the requirements? So first thing is you are actually requiring the trizol to require the 1.5 ml Eppendorf, you require the centrifuge, you require the chloroform, isopropanol, RNA 3 water, micropipettes and the tips and the test specimen.

So you require the either the tissue or the cell. So first step is that you are going to either take the tissue or the cell culture cells and then you are going to do the homogenization or the lysis. And once you do the vertexing and all other kind of things, then you are actually going to add the trizol and that actually is going to have the phase separations. So you are going to have the aqueous phase and then you are also going to have the organic phase. In the organic phase, you are going to have the lipid and you are going to have the protein, right. And then you collect the aqueous phase and then you add the ethanol and that is how it is going to form the pellet, the RNA pellet that you do by air dry and then you add the RNA 3 water and resolve things.

So in the step one, you are going to add the trizol reagent to the cell and incubate at room temperature for five minutes. Then you transfer the cell lysate to a 1.5 ml centrifuge tube and add 0.

2 ml of chloroform. So this is what you are going to do, right. In the step one, you are going to add the chloroform, mix it thoroughly and incubate at room temperature for five minutes. Then you centrifuge the mixture to the centrifugation at 12000 Rg for 15 minutes at 4 degree. Transfer the aqueous phase containing the total RNA to a fresh tube and precipitate the RNA by adding the 0.5 ml of isopropanol followed by incubation at room temperature for five minutes, 10 minutes. Then you centrifuge the precipitate at 12000 G for 10 minutes at 4 degree Celsius.

And then you discard the supernatant and air dry the RNA pellet for 10 minutes and resuspend in 20 microliter of RNA 3 water. Remember that this is very important and you can actually be able to either purchase the RNA 3 water from the commercial vendors or you actually can prepare the RNA 3 water in a laboratory. So it is not very difficult part. Then you perform the agarose gel electrophoresis to check the integrity of the RNA. This anyway we are going to discuss when we are going to discuss about the northern blotting.

So that time we are going to discuss about how you can be able to run the RNA gels and how you can be able to test whether the RNA quality is good or not. The RNA isolation by the tribal method is showing after adding the chloroform and centrifugation the mixture separate into three phases with the upper phase the aqueous phase containing the RNA, the interface containing the cell debris and the lower is the organic phase containing the protein and the liquid. So we have actually prepared a very small demo clips where we have actually going to show you how you can be able to isolate the RNA with the help of the tribal method and here the students have actually isolated the RNA from the bacterial cells. But you can actually be able to follow the similar steps even with the mammalian cells or the tissue. As I said you know when we are going to deal with the tissue you are actually going to homogenize the tissue so that you can get the single cell suspension.

So I hope this video or the demo video is going to be useful for you to advance your work. Today we will be learning about RNA isolations from the bacterial culture. As you can see this is a bacterial suspension already prepared. This is a suspension of the *Staphylococcus aureus* and we have already allocated.

So for RNA isolation we will be needing around 800 microliters. We have already allocated in the Eppendorf tubes. So to do that we need a laminar flow so that the contaminants does not get out. So as you already know we have already allocated around 800 microliters of bacterial suspension. So we will be today performing RNA isolation from this bacterial suspension. RNA isolation is a very tricky step as because it is easily degradable in the environment.

So to do that we have already given the UV for the whole hood. We have cleaned the pipettes with the 70% ethanol and all the tips and everything has been UV irradiated before use. So for RNA isolation our protocol is that RNA isolation can be done in from three simple steps. One is to homogenize the bacterial cells to take out the RNA from it, then to precipitate the RNA and then to purify the RNA. So first step we will do the homogenization of the RNA.

So to do that we will be using the tri reagent. This reagent is basically a trizol. So trizol contains basically gonadine, thiocyanate and phenol and it actually has an, it inhibits the RNA's activity so that the RNA is not degraded in the system. So now the protocol is we will be adding around 160 microliters of trizol in the suspension culture. The cap should be put down in the laminar as always and this trizol reagent will help to homogenize as well as protect the RNA integrity in the suspension. So to homogenize this, the thing is very simple we have to pipe it in and out faster so that the bacterial cells are homogenized as you can see in this step.

I am pipetting little bit vigorously in and out. After that thing is done we will be adding chloroform around one fifth of the total volume of this. So earlier it was 800 we have added around 160 more so it is around 950 microliters. So one fifth of the volume will be adding a chloroform in that. So chloroform will do one thing it will help separating out the phases in the mixture as such.

So now we will be adding chloroform in the trizol mixture we have already done. So chloroform we have allocated in this reagent bottle. So we will be adding around one fifth of the volume which is around 32 microliters. So as I told you before chloroform will help in separating out the phases. So actually trizol is very helpful by using trizol we can separate all the three components the DNA, RNA and protein as you will see in the subsequent experiment. So after adding the chloroform what we will do is we will tilt a little bit very gently which will not be harsh.

We will tilt a little bit and then we will leave it for incubation for 2 to 3 minutes at room temperature. So I will just put it here and wait for 2 to 3 minutes. So after that we will be putting it in the centrifuge. We have to centrifuge it for 12000 rpm for 4 degree for

15

minutes.

So that the phase separation might happen. So I have already put the in the centrifuge itself and I will start it now. So it will run for 15 minutes and then we will get back. So as you can see the run is just now going to stop. So we will be taking it out and then processing it further.

So after we take out we will see the layers getting formed. So after centrifugation as you can see very clearly that there are three types of layers which are formed. This is an aqueous phase which has the RNA. A white type of layer you see this contains DNA and the pink layer if you see that contains the protein. So from trizol we can isolate all three

DNA RNA and protein together.

But for today's experiment we will be doing RNA isolation. So we will be taking out the aqueous layer which is on the top. We have to be really careful not to take out the interface or the bottom layer. So we will be allocating this in the new centrifuge very very carefully. We are now allocating it the aqueous phase in the new Eppendorf tube.

So it would be you have to be very careful. So not actually you have to take it very slowly. Tilt a little bit and slowly pipette it out.

Don't just take the interface. That is the whole point. So we have taken the phase. We will be allocating it now. Still some is left so I will try to take more out of it. As you can see I am tilting it so that I will see clearly where the thing is going on. So it is always safe to not touch the interface.

So as you can see I have already allocated it and have not touched the interface. And this much RNA is more than enough for our experiment. Technically speaking we should not talk when we are doing RNA exhalation. As our when our aerosols from when we are talking might contain some type of RNAs which might degrade the RNA. So when we are doing RNA exhalation we should not talk much about it.

Now we will be adding isopropanol alcohol. This will help to precipitate the RNA from the aqueous phase. So how will we do that? We will be actually centrifuging it after adding equal volumes of isopropanol. So now we are centrifuging it for 10000 rpm for 10 minutes. So this will help out to precipitate the RNA after addition of isopropanol alcohol.

So now the centrifuge is almost complete. We will be moving to the last step of the RNA isolation which is the purification step by adding around 70% ethanol. And then again centrifuging it back. So now we have centrifuged it after the addition of isopropanol alcohol. Here as you can see the pellet is a little very less as well as it is a little bit shaky.

That might be due to the less suspension culture. I think the bacteria was not that much to give a very big thick pellet. But still the pellet is there. So now we have to very

carefully remove the isopropanol alcohol and add 70% alcohol which we have already prepared. So for the 70% alcohol preparation we have used mercury ethanol and we have used double distilled autoclave water.

Which is filtered using 0.2 micron meter membrane filters. So let us remove this isopropanol first. We have to remove very slowly so as to not take the pellet out. And if you are not sure whether the pellet will come or not you can leave it and then again take a small pipette. For example I am using a P1000 which is a grading of 100 to 1000. So for now I will leave this pipette and then I will take another pipette which is of a lower volume so that I can slowly take out the isopropanol.

So this is a pipette of 20 to 200 micrometers. So this would be a best fit for my thing now. So I will be setting it to around 100 micrometers. It is less volume and easily we can take it slowly. Now we have taken out the isopropanol.

And we will be adding 70% ethanol. So 70% ethanol we have to re-suspend the pellet. So what we have to do is we will add around 1 ml. So you can add 1 ml you can add 500 ml. This is a washing step also and purifying step also for RNA. So after you add you have to pipe it in and out a little bit so that the thing is re-suspended. And after that we will go ahead with the last centrifugation which is 10,000 rpm for 10 minutes at 4 degree centigrade.

So I am adding around 1 ml completely. And it is done. So we will go ahead with the centrifuge now. We are performing the last centrifugation step of the entire protocol. We will be doing 10,000 rpm for 10 minutes at 4 degree centigrade. And this will be helpful for washing the RNA. Now we have centrifuged the last step which was after addition of ethanol 70%. So the next step is to air dry it, take out the whole ethanol, air dry it and after air drying it for 10 to 15 minutes we will add RNA's water so that we can re-suspend the pellet in it.

So let's take out the ethanol. So after this step has been done, what we can do is because the pellet is there, we can just a little bit and let the residual ethanol come out here. Now we will keep it for air drying. Now it is after 15 minutes, the thing is fully dried. So now we will be adding the RNA's free water.

So this water we have already showed you before we have made the 70% ethanol. So we will be adding around 50 micrometers to it and then re-suspending the pellet again. So just add and re-suspend it a little bit.

Don't be too harsh. Pipe it a little bit and then just leave it. So the next step will be to

quantify this RNA. First we will quantify it using NanoDrop and secondly we will see the RNA using the gel electrophoresis. So the RNA can be stored at minus 20 degree centigrade. We should not store it at 4 degree for short time. So I will be storing this at minus 20 degree centigrade till further experiments are done in this.

Now today we will be doing RNA quantification. So this is an instrument, it is a NanoDrop machine. So this will measure the amount of RNA in one micrometer. So we will go to the nucleic acid section. As you can see in nucleic acid section there are lot of acids which can do it, can do protein and less everything.

So this is the whole template. Here you will see the concentration. This is 8 to 60 by 280 reading. A peak will come if the RNA is there. So not is set for dsDNA, means double stranded DNA. What we will do is we will set for RNA.

And now we have clicked RNA as you can see. We will proceed with the blank addition. To measure blank first what we will do as you can see here. So this is the portion where we load our sample area.

So we first clean it nicely both sides. Wipe it. This is our lint free wipe. After that we will add RNA's free water where we have already suspended in our RNA. So we will add one microliter exactly. So here addition is to be done on the red dot as you can see here.

So I will be adding directly here for RNA quantification. So now the whole sample has been loaded. I will close it and select blank as you can see here. So I will click on blank now.

So now it will adjust the blank reading first. And then after that we have to load our sample. So now the blank has set. We will load our sample after that. So to do that we have to first wipe out the RNA. So to do that we have to first wipe here again because we have already added.

And then what we have to do is take our RNA whatever we have made. Add it. And just add it. Close it and say sample. This green one is the sample. So we will just click it here. So it will measure the RNA content in the concentration of nanogram per microliter.

So as you can see RNA the concentration is 556.52. The A260 by 230 reading is 2.23 something which is a very good reading. There is no contamination. A peak, a nice peak we can see at 260. And so it is a very good amount of RNA without no contamination of

protein and DNA. So now I am sure you might have seen the demo video.

And this demo video could be very useful for you to replicate these steps in your laboratory. Although we have shown the steps with the help of the simple system like the prokaryotic system. But it can be replicated with the other type of cells also. So you can use the mammalian cells. You can use the yeast.

You can use the even the different types of tissue materials or you can actually be able to use the bacterial cells. The first step one is that where you are actually going to prepare the cell lysate. After that the subsequent steps are going to be remain identical whether it is a prokaryotic cell, whether it is a eukaryotic cell. Now once you isolated the RNA, you are actually going to have the first question. What will be the concentration of this RNA and whether the quality of this RNA is good or not.

Just remember that when we were talking about the DNA, we have also asked the same thing. So in the next step, we are going to talk about the DNA, RNA, whether the concentration and as well as the purity. So the purity of the RNA can be detected by the same way that RNA is also going to absorb very strongly at 260 nanometer. So what you can do is you can take the absorbance at the 260 nanometer. And once you do the RNA at 260 nanometer, it is actually going to give you the values. So if there is a, you know, the pure RNA, right, if it is a pure RNA, it is actually going to have the very specific absorbance at the 260 nanometer.

If it having the protein contamination, then it is actually going to have the ratio of 260 by 280 vary from the RNA species. So then it is actually going to have the, so if it is a pure RNA, the 260 by 280 would be around 2, right. Because RNA absorbs very strongly at 260 nanometer. But if it is having the protein contamination, then it is actually going to have the level at less than 2, right.

And that is how you can be able to know the purity of the system, right. Or you can be able to have the purity of RNA. Now the next question comes is how you can be able to do the estimations. So the estimation can be done at the absorbance, right. So you can take the absorbance at 260 nanometer and you can be able to use the formula to determine the excitations or you can actually be able to use the colloidal chip method. So RNA estimation by the Orsenol method.

Remember that when we were talking about the DNA estimations, we have talking about the, we have discussed about the DPA method, right. So what is the principle? This is a general reaction for pentoses and depends upon the formation of furfural when the pentose is heated with the concentrated hydrochloric acid. Orsenol react with the

furfural in the presence of the ferric chloride as a catalyst to give the green color which can be measured at 660 nanometers.

So you are, what you are going to do is RNA. RNA is ribonucleic acid, right. Which means it is going to have the ribose as a sugar, right. Remember that for the DNA it is deoxyribose and that is why. So when you heat this RNA in the presence of the HCl, it is actually going to form the furfural, right.

And the furfural is actually going to react with the Orsenol. And it is going to give you the green color compound. Or solution. And that actually is going to have the lambda max at 665 nanometer. Okay. So this 665 nanometer. So what you can do is you can actually be able to do a calibration curve with the lambda absorbance at 665 nanometer versus the RNA, right.

So you can actually we take the different concentration of RNA and that is how it is actually going to give you the calibration curve. And then you can actually be able to run the same with the unknown samples. And then suppose this is the absorbance of the unknown samples and then you can actually be able to determine the concentration of the RNA. So what are the things you require? So you require the standard RNA solutions.

So you can require the 0.2 mg per ml in 1% per chloric acid or the buffer saline. Then you also require the Orsenol reagent. So you can dissolve the 0.1 gram of ferric chloride in 100 ml of concentrated HCl and add the 3.5 ml of 6% weight by volume Orsenol in the alcohol.

And then you also require the buffered saline. So you can actually make the NaCl and you can make the shifted buffer pH 7. So it is very simple. You do not require a lot of reagents also.

And then the procedure. So you can actually be able to pipette out the different concentration or different amount of the RNA like 0.2, 0.4, 0.6, 0.8 and 1 ml. And to a working standard to a release of labeled test tubes. Then you pipette out the 1 ml of the given sample in another test tube. Make up the volume to 1 ml in all the test tubes. Test tube with a 1 ml of distilled water serves as a blank, right? So that is the blank reaction so that you know what will be the background absorbance of the Orsenol reagent itself.

Then you add the 2 ml of Orsenol reagent to all the test tube including the test tube labeled as the blank and as well as the unknowns. Mix the content of the tube by vertexing or shaking of the tube and heat on a boiling water bath for 20 minutes. So this step and then cool the content and record the absorbent at 660 nanometer against the

black, right? Then plot the standard curve by taking the absorbance concentration of the RNA along the x axis and the absorbance at 650 via the y axis. Then from this standard curve calculate the concentration of the RNA in the given sample.

So this is the table what you are going to use. So from the standard RNA stock you are going to have the 0, 0.2, 0.4, 0.6, 0.8 and 0.1. Then you add the water so make up the volume at 1 ml. So total volume of the RNA and the water is going to be 1 ml.

So for example in this case you do not have the RNA so you are going to only take the 1 ml water. In this case you have 0.8, 0.2 so you are going to add 0.8. So total is actually going to be 1 ml, okay? And then take this table and then for unknown you are going to take the 1 ml of unknown. You can take the other values of unknown also but accordingly you are going to add the water. And then you are going to add the volume of the reagents. You are going to add the 2 ml of the Orsenol reagent which you are going to prepare with the help of the ferric chloride and HCl and Orsenol.

And then you incubate this in boiling water bath for 20 minutes then you cool down and then you are going to take the absorbance. So absorbance what you are going to get for the 0 RNA is going to be treat as the blank, right? And this has to be subtracted although you are not going to get the 0 values but you are going to subtract that value so it is going to be 0. And then you are going to have some values for other value, other concentrations and for the unknown as well. Using these values you can be able to draw a calibration curve. So what you are going to do is you are going to have the absorbance at the 665 versus the RNA concentrations, right? Which means the microgram.

And then you are going to have the standard curve. So with the help of the standard curve you can be able to determine the concentration of the unknown RNA species, right? And if you want to read more about the RNA estimations and other kinds of things you can be able to go with this Plummer's book. This is a very important, very interesting book which is dealing with the practical aspect of the biochemistry. And it is very interesting because it gives you the step by step, you know the steps how you can be able to follow and how you can be able to prepare the recipes and so on. So this is all about the RNA what we have discussed.

What we have discussed? We have discussed about the structure of the messenger RNA and the different parts of the messenger RNA and what are their functions. And then we have also discussed about how you can be able to isolate the RNA and how you can be able to verify the RNA with the help of the estimations and with the help of the purity of the RNA. So with this I would like to conclude our lecture here. Subsequent lecture we are going to discuss about the proteins which are going to be the building block. And

then we are also going to discuss about the different types of enzymes which are actually going to be participate into the molecular biology. So with this I would like to conclude my lecture here. Thank you. .

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 03
Basics of Biomolecules
Lecture-14 Amino acids

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati and what we were discussing, we were discussing about the biomolecules. In today's lecture, we are going to start discussing about the new biomolecule and the new biomolecule what we are planning to discuss is the proteins. So as you can see that the proteins are actually being considered as the building blocks and these protein molecules are required for maintaining the, you know, the whatever the losses we have right. Suppose there is a loss of protein and we require the protein to replenish or if they will be a damage, then also we require a protein to reconstruct the damaged portion. Apart from that the proteins are also been responsible or also been functional as the enzyme.

So and that these enzymes are required for running the different types of metabolic pathways. I am sure we have already discussed a lot about the role of these enzymes into the carbohydrate metabolism when we were discussing about the glycolysis and the shape cycle. So subsequent to that, let us start discussing about the protein molecules. So as we said, you know, protein is a biomolecule and it is required for the as a building block.

So it is required at the building block, then it is required as an enzyme and then it is also required for running the metabolic reactions. Proteins are even responsible for making the different types of structures in the body, such as like we could be able to stand and you know walk because of the fibers which are made up of the protein. So all these we are going to discuss in this particular module. So let us discussing about the protein molecules. So protein are the polymer of the amino acids and they are joined by a covalent bond which is known as the peptide bond.

Each protein can be broken into the constituent amino acid by a variety of method to study the free amino acids. So what you can see here is that this is a generalized structure of amino acids, what you will see is you have a carbon in the center which is called as the C alpha carbon and then this C alpha carbon is attached to the four groups you have amino group, you have a carboxyl group and then you have the H and on the fourth side we have a side chain. So depending upon the side chain, we can have the different types of amino acids. So there are 20 different types of amino acids which are

present in the proteins and these amino acids are being joined by a covalent bond known as the peptide bond. So what is the structure of the amino acids? So amino acids share the common structures, all the 20 amino acids are the alpha amino acid with the common structures.

Each amino acid has a carboxyl group and the amine group attached to the primary carbon which is called as the alpha carbon, they differ from each other in terms of the side chain or the R group. The side chain varies in the structure, chemical nature and that influence the overall property of that particular amino acid. So what you see is this is the generalized structure of the amino acid where we have the C alpha carbon and that C alpha carbon is attached to one side as the amino group, the other side is called as the carboxyl group and that is why the name suggests is the amino acid. So amino acid means the compound which will have the amino group on one side and the acid group on to the other side and in between it is going to be connected to the H as well as the functional group which means the R group. So depending on the R group, the amino acids could be of different types.

The side chains actually varies in terms of the structures, chemical nature and that has influence on the overall property of the amino acid. Each carbon is attached to the four different groups. This is what we have discussed already. This is the C alpha carbon and it is connected to the four different groups making it chiral center to give the stereoisomers because this C alpha carbon is connected to the four different functional groups like the functional group number one which is the amino group, functional number group number two which is the acid group and then it has the hydrogen which is the third group and then it has a side chain or the R group which is the fourth side. So because of this four different groups attached to this particular carbon, this carbon is actually is a chiral center.

So it actually gives the stereoisomers. There are two common forms of the stereoisomers called as the enantiomers found in the amino acid. These are the non superimposable mirror image to each other. So this is what you see here is actually a enantiomer. So what you see here is actually a mirror image.

So if you see all the groups what is present on this side is present on to the this side in this particular amino acid. This is the example of the L-amino acid versus the D-amino acids. So compared to these all the amino acids only the glycine is the only amino acid which does not have the chiral center because glycine the R chain is H. So that is why the glycine does not show any kind of the chiral center. So amino acids are classified by the R group as this different amino acids are classified based on the side chain or to the R group.

These 20 amino acids are denoted by the first letter or the three letter codes. So these are the few amino acid what you see here. So we have the 20 different types of amino acid depending upon the side chains. For example we have the arginine, glutamine, phenylalanine, tyrosine, tryptophan and so on. So we have the glycine glycine and all that.

And all these amino acids are denoted by the two conventions. One is the single letter conventions or the third and the second is the three letter conventions. So this is very easy to remember. So even how the people are you can actually be able to remember the single letter or the three letter convention is for the single letter you have to just go with the first alphabet of that particular amino acids. For example in the case of alanine.

So alanine the first alphabet is A. So I can say that the alanine single letter code is A. Whereas if I want to go with the three letter codes I can just take the initial three alphabets. So then the three letter code is actually going to be alanine. So the same way you can actually have that other amino acids also like for example you can have the leucine.

So for the leucine if I have to remember the single letter code it is going to be L and if the three letter codes then it is going to be leu. This is not true for all the amino acid because you can also have the amino acids which is also be starting with the A alphabets for example we have the arginine. So if we have the arginine then it is for the these kind of amino acid you cannot take the A because A is already been alerted for the alanine. So what we can do is we can just go with the sound what is coming from this amino acid. So what I mean is coming arginine.

So when you speak the arginine the sound comes as R. So you can actually give the single letter code to R. So that is why these are the exceptions. These are the things what you have to remember. For example the arginine you have to remember that the single letter code for arginine is R and the triple letter code is A R G.

Similarly we have already been alerted the L to the leucine. So you can actually cannot alert the L for the lysine for the lysine the single letter code is K and thus the triple letter code is L by S. Okay the same is true for the other amino acid. For example the glycine you can easily take the single letter code as G and the triple letter code as GL by. So for the few amino acids you might have to remember the single letter code or the triple letter code but all other amino acids you are just going to follow the convention that the first letter of that particular amino acid you are going to take as the single letter code.

The name of the amino acids were trivial or the classical in few cases derived from the food source from which they are been isolated the first. This is a scientific way of giving the name of or the single letter or the triple letter code for these amino acids. But when the people have started discovering these amino acids they have given the name based on the food sources from which they have been isolated for the first time. For example the asparagine asparagine was isolated from the asparagus. And the glutamate is isolated from the wheat gluten whereas the tyrosine is isolated from the cheese and in the Greek the tyrosine is in the Greek the tyros means the cheese that is why the name of the amino acid is tyrosine.

Similarly the glycine, glycine is derived for its name due to its sweet taste because in the grief you have the glycos which is called as the sweet. So these are the conventional name which are been given but later on people have started giving the name based on the systematic scientific way of doing it and that is how they have given these kind of names. As I said you know remembering these single letter code or the triple letter code is very important because when we are going to write the sequence of the amino acids or for a particular protein you cannot write the full name. And that is why the single letter or the triple letter codes are being used in that particular cases. Now the amino acids are as I said you know amino acids are classified based on the R group so we have the 20 different amino acids and all these amino acids are varying in terms of the different types of properties.

So what you see here is I have given you a composite table where I have given you the three letter codes or the single letter code. Then as far as the molecular weight is concerned the molecular weight is also depending on the side chain as well as the other groups. So you can see that the molecular weight is 89 in the case of alanine whereas its molecular weight is 204 in the case of tiptofan. So it actually varies a lot but on a generalized way the amino acid molecular weight is considered to be 110 Dalton. So if I say there will be a question that what will be the how many amino acids are present in a

20 KDA protein.

So if this is a generalized question what you can do is the number of amino acid if you want to calculate then what I will do is I will just take the 20,000 KDA 20,000 Dalton and I will divide that by 110 and then what you are going to get is you are going to get the number of amino acid. So that is a generalized term if I say you should calculate the number of amino acids present in the actin protein then and if I give you a sequence of that amino acid or the sequence of that particular protein then the or if I say I give you a sequence of the amino actin protein or you give me the molecular weight then in that case the situation is going to be different then what you have to do is you have to first count the number of amino acids. For example, if I say number of glycines so suppose

the number of glycines are 10 for example so you can just go with the glycine table and then you say glycines molecular weight is 75 so 75 into 10 is 750 Dalton and that is how you are actually going to do for the calculations you are going to calculate the number of other amino acids like how many aspartates are present how many arginines are present how many lysines are present how many tryptophan are present and then you just keep multiplying that and if you add all those numbers then it is actually going to tell you that what is the molecular weight of the actin proteins. So, this is just a generalized term. So, this is the molecular formula what is being given for each and every amino acid then it is the residue formula what is given and then you have the residue weight which is may like you if you remove the water molecule then this is going to be the molecular weight then it is also giving you the pKa values and if the amino acid has two functional group then it is also going to give you the pKb as well and it is also going to give you the pI value.

So, these pKa values are actually going to calculate used to calculate the charge on that particular amino acids whereas the pI is actually going to calculate the charge. So, it is going to let you to calculate the charge of that particular amino acids at that particular pH. So, in a particular pH so, you know that at that particular pH the amino acid is going to be neutral. So, now the amino acids are classified by the R group. So, you can have the different types of R groups you can have the nonpolar aliphatic R groups.

So, the R group in this amino acids are nonpolar and the hydrophobic examples includes are the alanine, valine, leucine, isolecine, glycine, methionine and proline. Then we have the aromatic R groups which means these are the groups where you going to have a benzene ring. So, these groups are going to be contained the benzene ring whereas in this case it is not going to contain the benzene ring it is going to be a linear chains. So, the R group in these groups are hydrophobic side chains for example, the phalanine, tyrosine and the tryptophan. So, you might have to remember the structure of these amino acids if you want to understand the functional role or the crucial role going to be played by these amino acids.

Then we have the polar uncharged groups which means these amino acids are not going to have the charge which means they are not going to be positive or negative charges, but they are polar in nature. So, the R group in these amino acids are uncharged and they are more polar than the hydrophobic amino acids example includes are serine, threonine, cysteine, asparagine and the glutamine. Then we have the negatively charged R groups which means they are actually going to be called as the acidic amino acid because they are actually going to impart the negative charge into the protein. So, the R group in these amino acids are acidic with the net negative charge examples are the aspartate and the glutamate. Then we have the positively charged R groups and these amino acids are

called as the basic amino acids.

They are actually going to give you the positive charges and R group in these amino acids are basic with the net positive charges examples includes are the arginine and the lysine. Now when you take a protein and you want to calculate you want to know what will be the amino acids are present what you can do is you can just do acid hydrolysis and if you do the acid hydrolysis is actually going to give you the mixture of amino acids. Now once you have the mixture of amino acid you can resolve these amino acids on to a thin layer chromatography. So, thin layer chromatography is a technique which actually is going to give you the spot for the each and every individual amino acids and depending on the intensity of that particular amino acid and depending on the number of spots what you are going to get from this particular protein you can be able to calculate the number of amino acid as well as the type of amino acids present in that particular mixture. So, that you are going to do when you are going to analyze the amino acids.

So, how you are going to do the analysis of the amino acids? The thin layer chromatography technique is an analytical chromatography to separate and analyze the amino acids from the protein. In this method the silica or the alumina as a stationary factor a stationary phase is coated onto a glass or the aluminum foil as a thin layer and then a sample is allowed to run in the presence of the mobile phase. In comparison to other chromatography technique the mobile phase runs from the bottom to top from the by the diffusion whereas in most of the chromatography technique the mobile phase runs from the top to bottom by the gravity. So, what you are going to do is you are going to take a thin aluminum foil and on or the glass plates and then this aluminum foil you are actually going to make a thin film of the alumina or the silica and then what you are going to do is you are going to apply your amino acid which you want to dissolve and then you are going to keep this into a solvent system. So, once the solvent is going to run in from the bottom to top it is actually going to take up the your amino acid also along with that and then it is actually going to give you the spot as per the and it is going to give you the mixture of the amino acids and that is how it is actually going to dissolve.

Now what you see here is in the TLC you are going to have the moment of the solvent from the bottom to top because of the diffusion compared to that in a conventional chromatography you are going to see always running of the mobile phase from the top to bottom. Like example if you run a column right the column you run from the water from top to bottom whereas in the case of thin layer chromatography the solvent runs from the top to bottom because of the diffusion. When a sample runs along with the mobile phase it gets distributed into the solvent phase and the stationary phase right when it runs it is actually going to be you know going to get distributed so either it will go with the solvent fronts right so you can see that all the several molecules are running right. So

either so if this is the molecule either it will go along with the solvent front or it will actually go to remain with the silica front right and because of that it is actually going to be distributed because that kind of affinity and that kind of differential behavior is going to be different for the different amino acid molecules. The interaction of the sample with the stationary phase retard the movement of the molecule whereas the mobile phase implies and as the effective force.

So what will happen is that when you are going to apply the sample onto this it is actually going to experience for example this spot right. So if you have a spot here it is going to experience two different types of forces it is going to have a driving force right. So driving force is actually going to be exerted by the solvent molecule. So solvent molecule is going to try to push this molecule towards the top side because it is running right but whereas it is actually going to make the interaction with the these silica particles which are present on the this TLC plate and because of that there will be a retardation forces. So that retardation forces is going to be by the matrix molecule and because of that it is going to have the two opposite forces one which is going to be on towards the top side and the other one is going to be onto the bottom side.

So it is going to be distributed so and it is going to be immobilized onto the plate. Now suppose the force caused by the mobile phase is F_m and the retardation force by the stationary phase is F_s . So then the effective force on the molecule will be $F_m - F_s$. So this is what I am going to I was trying to explain you right. So if this is the spot on this spot you are going to have the upward forces which is called as F_m or the mobile force by the mobile phase whereas it is actually going to have the retardation forces which is going to be caused by the stationary phase which is called as the F_s .

So this molecule is actually going to run effectively by a force which is going to be $F_m - F_s$. Now the $F_m - F_s$ so where this molecule is going to be stopped right it is going to stop the place where the F_m is going to be equivalent to F_s right and that $F_m - F_s$ is going to be different for the different molecule and that is why they are actually going to be present at the different places. For example for this molecule the F_m is too big for the F_s which means it is still having a effective charge at this point and that is why this molecule runs for the longer period of time but at this point when it got immobilized again for this one also the F_m is equivalent to F_s . So the molecule is immobilized on the silica gel where the F_m is going to be equivalent to the F_s and the position is controlled by the multiple factors. Nature or the functional group what is present on to the molecule or the analyte.

So if the molecule is going to interact with or suppose it has the functional group and it is going to interact with the silica particle then it is actually going to increase the F_s . If it

is going to increase the F_s then the $F_m - F_s$ is going to be small right and if this is the case then it is actually going to be immobilized towards the spotting points. Then it is also going to depend on the nature on the composition of the mobile phase. So depending on the mobile phase also it is actually going to be different right if the mobile if the molecule is very soluble into the mobile phase then the F_m the value of F_m forces are actually going to go up right and that is why it is actually going to run very far away from the spotting place right. Then it also depends on to the thickness of the stationary phase because that also is going to have the effect on to the retardation forces and then it also has the functional group.

So functional group what is present on to the stationary phase. So apart from the functional group what is present on to the analyte molecule the functional group if the functional groups are also present on to the silica particles right. For example if you take the silica particle or if you take the functional silica particle they may have the higher affinity for this particular molecule and as a result the F_s will actually go up and if the F_s will go up it is going the $F_m - F_s$ is going to be small and that is why it is actually going to immobilized very soon and it is going to be closer to the spotting points. So you can imagine that if I want to see the differences like what I can do is or if I want to know where the molecule is going to immobilize what I can do is once it got immobilized I can take a distance from this molecule to the origin.

So this is the origin point. So for example at this point I have started putting the spot of the mixture and then the solvent started running right. So when the solvent reached to the end of the plate or it reaches to at least the 75% then what I can do is I can just stop this I can develop this spot and then I will calculate the distances what is run by the solvent and I can also calculate the distances run by the individual spot. So for example this is the spot number 1 this is the spot number 2 this is the spot number 3 and the distance of the spot number 1 is d_1 plus distance of the spot number 2 is d_2 and the distance of the spot number 3 is d_3 . So what I can do is I can just calculate the R_f values and R_f value what is the formula is that the distance of the analyte right distance of the analyte spot which is like d in this case so d_x divided by the distance of the solvent. So distance of solvent like the in this case d_s okay and this is going to be fractional.

So the maximum R_f what is possible is 1 and it going to be the fraction of 1 right because this is the ratio right and this R_f value is going to be dependent on to the solvent system right depending on the solvent what you have taken depending on the matrix material and depending on the conditions in which it is run. So if you are going to maintain all the 3 constant right if you maintain the same solvent system if you maintain the same matrix and if you run it under the identical conditions the R_f value is not going to be changed even if you run it for for example if I run it for 50 centimeter it is going to

be distributed accordingly. So that for example if I have Rf value of 0.5 okay. So if I run it for 50 centimeter the spot is going to be formed as 20 centimeter 25 centimeter if I run it or 100 centimeter right then the spot is going to be formed at 50 centimeter because it is going to maintain the ratio of Rf is equal to 0.

5. So if Rf is 0.5 which means the ratio of dx to ds is going to be 0.5. So that does not depends on the how much length you are going to run the TLC plates it is going to be always be immobilized at the 50 percent distance right if the Rf value is 0.5. So this Rf value is constant and that side the Rf value can be used to characterize the different types of amino acids.

Now the question comes how you can be able to determine the Rf values you can actually be able to run the TLC plate that you can be able to run the thin layer chromatography and that is how you can be able to calculate. So how you can actually be able to run the thin layer chromatography. So these steps are required to perform a thin layer chromatography to analyze the complex samples these preparative and operational steps are as follows. So in the step number one you are actually going to make the spotting but before that you are actually first going to take a thin layer plate.

So what you see here is this is the thin TLC plate right. So what you are going to do is what you take the TLC plate you cut the TLC plate as per the number of sample what you are going to place right on the width wise and its height is also going to be as per the chamber of your TLC plate right. So you are going to run it in a chamber right and then what you are going to do is you are going to take a scale and you are going to put a line at this line should be above to the solvent front right because it is going to dip right ultimately you are going to dip this. So then you put a line and then on this line you are actually going to put the spots. So the line is drawn with a pencil little away from the bottom the sample is taken into the capillary tube or in a pipette. So what you can do is just take the sample into a capillary tube the capillary is touched onto the silica plate and sample is allowed to dispense.

So what will happen is when you touch the silica when you touch the capillary to the that particular silica it is actually going to suck the sample automatically by the action of diffusion right. So it is important that the depending on the thickness of the layer a suitable volume should be applied spot is allowed to dry in air or a hairdryer can be used instead. Then we have the running of the TLC. So once a spot is dried it is placed in the TLC chamber in such a way that the spot should not be below to the solvent level, solvent level front is allowed to move until the end of the plate.

So what you can going to do is just keep it into a TLC chamber. So you can actually be

able to develop a TLC chamber into a beaker or into a thin small chambers depending on the type of the solvent right. So then you have to cover this with a some you know through the lid right rather you can just put a cap right and why it is important because so that the solvent what you have put is actually going to form the vapor and that vapor should be condensed otherwise it is not going to give you a upward moment. After the so then you are going to place the TLC into this right. So you are going to plate the TLC plate and make sure that this line should be above to the solvent front so that it is not going to get dissolved into this solvent before running right. The analysis of the chromatography plate, the plate is taken out from the chamber and air dried if the compound is colored it forms the spot and these substances there is a no additional staining required.

There are two method of developing the chromatogram. So what you can do is later on you can just take out this plate and then you air dry. So if the analytes like the amino acids are colored then actually they are going to give you a spot and then you can directly take the all the sort of measurements so you can actually know that this is the like the ds right. So this is the solvent what you have run and then you can just take the calculation of this and it is actually going to be dx and then you can actually be able to calculate the RF value by dx by the ds . So for this the position of the spot is very important right. So if it is colored compound there is no need to have the any kind of additional staining procedures or any kind of procedures you can actually be able to do this right.

But if it is not then you have to go with the staining procedures. In the staining procedure the TLC plate is sprayed with a staining reagent to stain the functional group what is present into the compound. For example the ninhydrin is used to stain the amino acids. So if it is not then you can actually be able to use the staining procedures. If you want you can actually go with the non staining procedure as well.

So you can use the non staining procedure. In the non staining procedure spot can be identified by the following method. You can use the autoradiography a TLC plate can be placed along with the x-ray foam for 48 to 72 hours. Exposure time depends on the time and the concentration of the radioactivity and then the x-ray film is processed. So you can actually have the radioactive amino acids and that is actually going to be exposed to the x-ray film and then it is actually going to give you the spot onto the x-ray film and then you can do all the calculations from this particular spot. Like for example you can do the RF calculations you can actually do the dx and ds and that is why you are actually going to get the RF value and that RF value can use for identification of that particular unknown spots.

The second is you can do the fluorescence. So several heterocyclic compound gives the fluorescence in UV due to the presence of conjugate double bond system. TLC plate can be visualized in the UV chamber to identify the spot. So what you see here is this is the typical UV chamber where you have the UV bulbs and this is the chamber right this is the lid of that particular chamber. So what you can do is just open from here and place the plates under this and then you can just turn on the UV lights. So you can have the two different types of UV lights which you can use as per the wavelength and then from this size because the UV light is dangerous for the eye.

So that is why you can actually be able to observe not directly but through this particular observing window and what you see here is that this all the spots are visible right. So under the UV and then you can actually be able to use the camera or some other acquisition system and that is how you can actually be able to capture the image. Apart from that you can also use the iodine. So you can also use the iodine staining right. So you can actually be able to incubate the TLC plate into the iodine chamber and this iodine chamber is actually going to stain the spots as well.

So the proteins are as we discussed that proteins are the polymer of the amino acid they are joined by the covalent bond known as the peptide bond. A peptide bond is formed between the carboxyl group of the first and the amino group of the second amino acid with the release of the water molecule. So this is you will see that this is the amino acid number 1 and this is the amino acid number 2. So when they will go by with the condensation reactions so what will happen is that the OH of this acid and the H from this is actually going to combine and that is why there will be a bond which is going to be formed between the two amino acids and that is how there will be a loss of water and this is a dehydration reaction.

So it is a dehydration synthesis or the condensation reactions. The peptide bond has partial double bond character due to the resonance and the CN bond is not free to rotate. But the bond between the N to alpha like to N to the alpha and the C to C alpha can be able to rotate through a dry handle angle designated by the phi and psi. So what you can do is this peptide bond is rigid it is not allowed to rotate but the bond between the N to alpha and the C to C alpha can be able to rotate through the dry handle angle which are designated as the phi and psi and these angle can be able to rotate from the minus 180 to plus 180 with the few restrictions. To exploit this particular type of phenomena, the Indian scientist G N Ramchandran has determined the possible phi and psi angles for a particular amino acid by synthesizing the tripeptide with the amino acid of interest in the middle.

So what he has done is he has actually synthesized a tripeptide. So for example, if he

wants to calculate the phi and psi angle for A he has made a tripeptide with C and B. So you can keep changing these tripeptide and that is how you can be able to calculate under different conditions what will be the different psi phi angles are possible and that is how you can be able to make a map between the psi versus phi. So you can actually be able to make a angle between the map between the psi and phi and then you can calculate that under different combinations how much these psi and phi angle are going to vary for the A molecule and that is how you can say that okay A will go from this psi angle to this phi angle. And based on that he has actually developed a map which is called as the Ramchandran plot and that Ramchandran plot is used to define the region of the allowed rotation for the amino acid present in a protein structure and he what he proposed is that he you can use this particular type of plot to say whether a solved protein structure is correct or wrong because if it is incorrect then the psi and phi angle are not going to be present in that particular defined regions. So that is what you see here is that you are actually going to see the different regions and it is going to be what you see here is the plot between the between the psi and phi and that shows the location of the different types of structures what is present in the protein structures and so on.

And that is how it is actually going to give you the distribution of that particular amino acid in the protein structures and how much its phi and psi angle are going to vary. So with this brief discussion about the Ramchandran plot and as well as the brief discussion about the amino acid I would like to conclude my lecture here. In our subsequent lecture we are going to discuss about the some more properties structural properties of the proteins. So what we have discussed so far we have discussed about the biochemical properties of the amino acids and we have also discussed in detail about how you can be able to analyze the amino acid using the thin layer chromatography. So with this I would like to conclude my lecture here. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 03
Basics of Biomolecules
Lecture-15 Protein

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati. And what we were discussing we were discussing about the biomolecules and in the previous lecture we have discussed about the proteins. So we have started discussing about the proteins that the proteins are made up of the amino acids and these amino acids are having a general structures where they have the central C alpha carbon and on this central C alpha carbon we have the four different types of functional groups which are attached to each on one side it has the amino group on the other side it has the carboxyl group and third side it has the hydrogen and the fourth side it has the functional side chains and based on the functional side chains the amino acid can be classified into four different groups 20 different types. It could be vary from the very simple glycine to a very, very complicated tryptophan and depending on the side chain, it can be of the different molecular weights and different types of properties. So, it could be a hydrophobic amino acids, it could be a hydrophilic amino acids, it could be a polar amino acids, it could be negatively charged amino acid or it could be a positively charged amino acid.

And in addition to that we have also discussed in detail about the thin layer chromatography and how the thin layer chromatography can be used to analyze the different types of amino acids. So, in today's lecture we are going to discuss more about the protein structures. So, let us start discussing about the protein structures. So, as we said you know protein is made up of the 20 naturally occurring amino acids.

A typical amino acid contain amino group and a carboxyl group attached to the central alpha carbon. The side chain attached to the central alpha carbon determine the chemical nature of the different amino acids. So, what you see here is the you have the C alpha carbon and that C alpha carbon is attached to the 4 different types of functional groups and these 4 different types of functional groups and depending on the R side chains, you can have the 20 different types of amino acids. Peptide bond is connecting the 2 individual amino acids and that is how they are actually giving the polypeptide chain. Each amino acid is linked to the neighboring amino acid through an acid amide bond between the carboxyl group and the amino group of the next amino acids.

Every polypeptide chain has the free amino group and the free C terminal groups. That

is why the primary structure of a protein is defined as the amino acid sequence from the end to the C terminals with a length of the several hundred amino acids. So, these are the 20 different amino acids, but we have already discussed in the previous lectures. So, the primary structure is defined as the amino acid sequence from the N terminals to the C terminals. So, do you see that it starts with the N terminals.

So, this is the N terminals and then it ends up with making a combination with the all of the amino acid to the peptide bond, but ultimately with the last amino acid it is going to have the carboxyl group which is going to be free. And that is why the protein is going to have the amino group and then it is going to have all the amino acids and then it is going to have the carboxyl group. So, that is why the protein has the two ends one is N terminals and the other one is called as the C terminals. The ordered folding of the polypeptide chain give rise to the 3D conformation known as the secondary structure of the protein such as the helix sheets and loops. So, this is what you see here is the primary structures.

So, all the protein all the amino acids are present and the first amino acid is actually going to have the N terminals whereas the last amino acid is going to have the C terminals. When this primary structure is getting folded, it is actually going to give you the secondary structures where you have the alpha helices, these are the alpha helices, then we have the beta sheets. So, you can see the beta sheets, these are the beta sheets and then it also has the turns. So, these are the turns what you see here. Arrangement of the secondary structure give rise to the tertiary structures, alpha helix and beta sheets are connected by the unstructured loop to arrange themselves in the protein structure and it allows the secondary structure to change their directions.

Tertiary structure defines the structure of a protein and the enzymatic activity or the nature of the structural protein. So, once the secondary structures are joined together by the loops or the turn, they are actually going to give you the tertiary structure. So, this is what you see here is the tertiary structures and if the protein has the multiple subunits, then it is actually going to give you the quaternary structure. For example, in this case, we have the subunit 1, 2, 3 and 4. So, all these 3 subunits, all 4 subunits are coming together and that is how you are going to have the quaternary structures.

So, different polypeptide chains are arranged to give the quaternary structure. So, depend that is why if you want to understand the protein structures, you have to understand all the 3, 4 different types of structures. So, we have the primary structures, we have the secondary structures, we have the tertiary structure and we have the quaternary structure. So, these are the different level of organizations, what is present into the protein structures. So, let us first start with the primary structures.

So, primary structures, the amino acid sequence of a protein is known as the primary structure, the order of the amino acid determine the folding of the protein to achieve the net minimum free energy and this is achieved in the multiple steps collectively known as the folding. So, if the primary structure is going to fold, so, this is what you see here is the primary structures where you have the length of the amino acids. Now, see here I am just showing you the single letter code of the different amino acid because it is almost impossible to write the full length or the full name of that particular amino acid because to save the space and these primary structures are actually going to fold to give you the secondary structures and that event is called as the folding. So, when the primary structure is going to fold into a proper three dimensional conformations, then it is going to give you the secondary structures. Now the question comes how we can be able to determine the primary structures.

So, if you want to determine the primary structures, you have to first achieve the primary structures and then you can be able to sequence the protein and you can be able to know the amino acid sequence of that particular proteins. So, these are the different steps what is being shown. So, what you can have is you can have the starting with the protein structure. So, this is the three dimensionally folded protein. So, here you are going to have the tertiary and secondary structures.

Now, what you have to first do is you have to first convert that into a linear chain of amino acid which means first you have to achieve the primary structures. So, it means you are actually going to unfold the protein by the chemical or the enzymatic method, then what you are going to do is because this length is going to be very large, then you are actually going to break the peptide or break the protein into small pieces. So, into small pieces and you are going to break the small pieces like this. So, they are going to be overlapping regions. So, that you can be able to add put them these things separately and then what you are going to do is you are actually going to label the terminides or the terminal amino acids.

So, you are going to have the labeling of the terminal amino acid in this way. So, you are going to have the labeling. So, labeling with the fluorescent dye in different and then you are going to do the sequencing, then you are going to identify that labeled amino acid that sequencing you can do by two methods you can use the Sanger sequencing method or you can use the Edman degradation method. So, this is what it is showing here. First you are going to start with the 3D fold structures, then you are actually going to

So, you are going to use the different types of treatments like you are going to use react

with the FD and B and all that. So, that it is actually going to break the disulfide linkages and then once the it is going to adopt the primary structures, then that you are going to degrade the primary structure with the help of the different types of chemicals or the enzymes and that is how you are going to get the small fragments. And once you got these small fragments, you can actually be able to do the sequencing and then once the sequencing is over, then you can be able to put these blocks together and that sequencing you can do either by the Edman degradation method or to the Sanger sequencing method. So, let us understand. So, in the state 1 this is the stage 1.

So, stage 1 you are going to convert the 3D conformations into the 1D or the primary structures. So, that you are going to achieve simply by the stage 1 where you are first going to break the disulfide linkages. So, stage 1 is the breaking of the disulfide bonds. So, you can imagine that this is the protein which has a disulfide linkages and the disulfide linkages interfere with the complete sequencing procedure as it does not allow the release of the cleaved amino acid from the peptide chain. There are 2 approaches to disrupt the disulfide linkages in a protein sequence.

In the first approach the protein is oxidized with a performic acid to produce the 2 cystic acid. So, why there is a need to break the disulfide linkages because if you do not break the disulfide linkages, even if this particular amino acid is actually going to be labeled and it is going to be hydrolyzed, it is not going to be released from the main chain because it is still having a it is bind to the main chain through a disulfide linkage and that is why it is important to break the disulfide linkages. So, disulfide linkages we have the 2 approaches in the approach 1 you can actually use the oxidation with the performic acid and that is when you do the performic acid treatment it is actually going to break the linkage between the disulfide linkages and that is how it is actually going to give you the 2 fragments. Whereas in the approach number 2, the protein is reduced by the DTT or beta mercaptoethanol to form the 2 cysteine followed by the treatment with the iodoacetate to form the carboxymethyl cysteine formation of the carboxymethyl cysteine is stopped the reformation of the disulfide bond. So, in the approach 2, what you are going to do is you are going to add the DTT.

So, DTT is a reducing agent. So, once you are going to reduce the disulfide linkages the S is S is actually going to get converted into SH and that is how it you are going to have the 2 peptide bonds where the disulfide bond is broken. But this has a problem because as soon as you have the SH and you have the reducing environment it is going to be remain as SH. But once it is actually going to be acquired the oxidizing environment again the S is going to be oxidized and again the SH is going to be get converted into the SS double bonds. So, to avoid that you are again going to react this with the carboxymethylene by the iodoacetate.

So, the in that case then what will happen is that the S is actually going to be tagged with this particular functional group. This is going to like a and that is how it is actually going to form the carboxymethylated cysteine residue and once you have this then they will not be able to come together even if the conditions are oxidizing in nature. Now when once this is done you can actually go back to the stage 2. In the stage 2 you are actually going to break the big polypeptide chain into the multiple fragments. So, in the stage 2, stage 2 is the cleavage of the polypeptide chain.

So, in this stage 2 is the cleavage of the polypeptide chain the protease and the chemical treatments are targeting protein have a specific recognition sequence and they cleave after a particular amino acid. So, this is one. So, stage 1 is this. So, stage 1 is over where you have actually destroyed the disulfide linkages by the two approaches what we have just discussed and now in the stage 2 you are actually going to cleave the protein with the enzymatic or the chemical methods. So, some of the common reagents what you are going to use for fragmenting the polypeptide chain is that you can use the enzyme.

So, you can if you use the trypsin enzyme the trypsin has the cutting side which is actually after the lysine or to the arginine which means wherever the lysine or the arginine is present for example, this is these are the peptide which are being generated by the trypsin. So, if you treat it with the trypsin it is actually going to cut wherever you have the lysine or the arginine. So for example, here it has cut here it has cut. So, the wherever you have the lysine and arginine it is actually going to cut and that is why it is actually going to generate the different types of fragments. Similarly, you can use the chymotrypsin.

So chymotrypsin is actually going to cleave the peptide after the phenylalanine tryptophan or tyrosine which means after the aromatic amino acids. So, then we can also use the pepsin. So pepsin is actually going to cleave the polypeptide chain after the leucine, phenylalanine, tryptophan or the tyrosine. And then you can also have the different types of chemicals for example, you can use the cyanogen bromide and cyanogen bromide is actually going to cleave the polypeptide after methionine. For example, these are the peptide sequences what has been generated by the cyanogen bromide.

So, you see this is the methionine and after the methionine it has been cut by the cyanogen bromide. Now, once you got these small fragments, then what you can do is you can take the individual fragments and then you can sequence these small fragments. So, once you got the sequence of these small fragments, then you can have to put them together and that is how you are going to get the sequence of the complete proteins.

Now in the stage 3, the stage 3 you are going to do the sequencing of the polypeptide chain. So, these are the you are going to have multiple polypeptide chains.

So, for the sequencing of the polypeptide chain, you can have the 2 methods one is you can use the Sanger's method or you can use the Edmund-Diggard-Dahn method. So, let us first discuss about the Sanger's method. So, once the polypeptide fragments are generated, we can start the sequencing of the each polypeptide chain, it has the following steps. So, first thing is you have to identify the N-terminus residues, the N-terminal amino acid analysis is being performed in the 3 step. Number 1, you are actually going to label the terminal amino acids.

So, as I said, you know, when we were talking about the primary structure, so primary structure has the N-terminal thing and then it has a C-terminal thing. So, what we are doing is we are first sequencing the protein from the N-terminus. So first amino acid we have to first, you know, do the sequencing from the N-terminus. So, for the first amino acid we are just using the labeling. So, we are labeling the terminal amino acids.

So, the chemical reaction is performed to label the terminal amino acid with the compound such as the Sanger's reagents like 1-fluorot-2,4-dinitrobenzene or D-PhenB and the Denzyl chloride. In most of the cases these reagents also label the free amino acids which are present on the basic amino acid such as lysine and arginine. Dinitrofluorobenzene reacts with the free amino group to form the dinitrophenyl amino acid complex. So, what we are going to do is you are going to take the D-FNB and then you if you add the D-PhenB to the first amino acid because it has the free amino group, it is actually going to and in the presence of the HF. So, there will be a release of the this group and then it is actually going to form a bond with the terminal amino groups and that is how the first amino acid R1 is actually going to be labeled.

Now, what the step 2 what you are going to do is in the step 2 you are going to hydrolyze the peptoene. So, that when you add the acid hydrolysis of the dinitrophenyl amino acid complex that lead to the breaking of the peptide bond to release the dinitrophenyl amino acid complex in the solutions. So, after this once the first amino acid is been labeled then you are going to do the acid hydrolysis and as when you do the acid hydrolysis it is actually going to break the bond between the first amino acid and the second amino acid and as a result the first amino acid which is already been labeled with the D-PhenB it is going to be released from the main chain. Then you are going to do the separation and the analysis of the derived amino acids. So A-HPLC or the TLC separation of the complex and comparing it with the standard amino acid is actually going to give you the name as well as the identity of this.

So, what you are going to do is once you got this amino acid then you can actually be able to run the TLC along with the standard TLC. So, what you can do is like for example, you can run a TLC like this. So, you can run all the 20 amino acids. So, you can actually make the all the 20 amino acid in the D-PhenB complex and then you can run and then you can also run the unknown sample. So, if you run the unknown sample it is suppose it goes the spot here and suppose this is the arginine.

So, if it goes to this, then you can say that the this is the arginine or you can actually be able to calculate the RF value of your unknown sample. And since you know the RF value of all the other amino acid complexes, you can be able to identify this. The other approach is that you can do the HPLC and you can be able to calculate the retention values. So, this is about the how you can be able to use the Sanger's method to sequence the proteins. Now, if you talk about the Edman degradation method, so in the Edman degradation method, it also has a similar kind of steps.

For example, the similar to the Sanger reagents, the reagents are different like where you are actually going to use the phenyl isothionate reacts with the terminal amino group to form a cyclized phenylthiocarbamoyl derivative. So, in this case, you are going to use the phenyl isothiocyanate and when it reacts with the terminal R1 group on to the peptide, then it is actually going to form a cyclized product. So, under the acidic conditions, the terminal amino group is actually going to be cleaved from the main chain as a thiazolidine derivatives and that is how you are actually going to have the first amino acid as the PTH. So, thiazoline derivative is extracted into the organic solvent and it forms the phenylthiodine amino acid PTH amino acid complex into the presence of acid. So, ultimately you are going to get the PTH complexes.

So, this is for the PTH complex of the R1 and then what you are going to do is you are going to run the PTH amino acid complex can be identified by the HPLC or TLC in comparison to the standard amino acid. Now once you have done this 1 to 4 for the first amino acid. So, when you do the first step 1 to 4 for the first amino acid, you can do the 1 to 4 again for the second amino acid because the first amino acid is been released. So, that is the remaining peptide chain is still there, you can use that and again do the another round of this. So, if you continue this like this, it is actually going to keep giving you the amino acid sequence from the N-terminus side.

So, if you step 1 to 4 can be repeated for the next amino acid in the polypeptide chain and that is how it is actually going to give you the whole sequence and that the sequence if you have the different fragments if you put them together it is actually going to give you the complete sequence of the all the protein peptides. Now as I said you know the protein is having the 2 chain types. So, you have the N-terminus side and then you have

the C-terminus side. So, we have just discussed the method like the Sanger's method or the Edmond degradation method to identify the N-terminus amino acids, but we can also do the sequencing from the C-terminus and that is how you can be able to identify the C-terminal residues. So, how we can do that the C-terminal residues not many methods are developed for the C-terminal amino acid analysis.

The most common method is to treat the protein with the carboxypeptidase to release the C-terminal amino acid and test the solution in a timely dependent manner. So, what you can do is you can just treat this with a carboxypeptidase. So, carboxypeptidase is a specific enzyme which actually releases the amino acid from the C-terminal side rather than the N-terminal side. So, if you use the carboxypeptidase it is actually going to release the amino acid and these amino acids are the C-terminal amino acid. So, once the amino acid is released you can identify that amino acid by the Sanger's or the Edmond degradation methods.

Then the stage 4 you are going to do the ordering of the peptide fragments. So, usage of the different peptide cleavage reagents produces the overlapping amino acid stretches and these stretches can be used to put the whole sequence. For example, when you generate the trypsin it is going to generate the first fragment like this, the second fragment like this, the third fragment like this. So, if you sequence this fragment, if you sequence the A fragment, if you sequence the B fragment, if you sequence the C fragment. So, what you see here is that A is having this portion which is overlapping, C is having this portion which is overlapping with the B.

So, by doing this overlapping sequencing you can be able to deduce the final sequence of the final length or you can be able to put them these fragments and that is how you can be able to do the sequencing. Then the stage 5 you can actually be able to locate the disulfide bonds. So, the peptide cleaved by the protein cleaved by the trypsin is performed with or without breaking the disulfide linkages. Amino acid sequence analysis of fragment will provide the side of the disulfide bond. Depressants of a disulfide bond will reduce the two fragments, will reduce two fragments and will appear as a single large fragments.

As we said, if there is a disulfide linkage present, it is not going to allow the release of the amino acid fragments. So, because of that if there is a suppose this is the disulfide linkages and so even if you cleave the this particular fragment. So, for example, if there is a disulfide linkage like this, if there is a disulfide linkage like this, if you cleave this with the protease, it is ideally should give you the two fragments. But if there is a disulfide linkage, it is still be going to bind like this and that is how it is actually going to give you the single fragments. So, if you get the single fragment, then there is a disulfide

linkage which is present.

Apart from these kind of methods and with the advancement of the mass data of different types of peptide fragments and all those kind of thing, the mass spectrometry method is also been used. So, in the recent past the mass spectrometry method in conjugation with the proteomics information is also been a popular tool to characterize the each fragment to deduce its amino acids. So, in this mass spectrometry data, what you are going to do is instead of doing this sequencing, what you can do is you can just simply calculate the peptide mass and that peptide mass actually because there is a complete database can be used to deduce the sequence. I have given you the reference. So, if you are interested more about reading the protein sequencing, you can actually be able to read through this particular reference.

Now, let us talk about the secondary structures. So, secondary structures the amino acid interact with each other and as a result the peptide chain folds into secondary structures. These secondary structures are the building blocks for the tertiary structure. So, these are the prime structure when they fold, they give you the secondary structures. Secondary structures could be of two types either it can be alpha helix or to the beta sheets. It is a helical structure termed as the alpha helix by the linear spalling.

In this structure, the polypeptide backbone is wound around a central axis with the R group of the amino acid protrudes outward from the helix bond band. In most of the protein the helix is right handed which means you are actually going to see the helix and in the helix the R groups are protruding outside. So, then we have the beta sheets. This is some more extended confirmation of the polypeptide chain where the R groups protrude from the zigzag structure in the opposite direction giving a alternate structure. Beta sheets could be the two types it could be the parallel beta sheets or the anti-parallel beta sheets.

So, it can be parallel or the anti parallel beta sheets. Which means either the beta sheets are running in the same direction, then it is called as the parallel beta sheets. If they are running in the opposite direction then it is called as the anti-parallel beta sheets. Then we have the turns. These secondary structures have no definite structures and they are present in the protein structure to change the direction of the running polypeptide. These are also found to places to connect the successive alpha helix and beta sheets, the number of amino acid and their preference in turn is not consistent.

The two protein can adopt the similar 3D conformation by changing the length and keeping the amino acid in the turn region of the structures. So, turn is actually a unstructured region and it is actually not having a definite structure, but it has a very

huge significance in terms of providing the flexibility of the different types of protein structures. So, for example, you can have the two different types of protein structures, their sequence, their amino acid sequence could be different, but they may adopt the identical structure simply by changing the length of the loops within the length. So, you can see that these are the loops. So, what you see here is this green color region is actually called as the loop or the turn.

Then we have the tertiary structures. So, tertiary structures, secondary structures forced to give the rise the higher order organizations commonly known as the tertiary structures and then the tertiary structures can still be packed and that is how you can have the quaternary structure. So, if the multipolypeptides are involved in the concentration of the protein, the tertiary structure of these different polypeptide chain come together to form the quaternary structures. Now, as we discussed about the methods to determine the primary structures, we have discussed about the Sanger's method and the Edmond degradation method. We also would like to discuss about the methods to determine the secondary or the tertiary structures. So, what are the methods to determine the secondary or to the tertiary structures? So, there are two approaches one is you can use the experimental methods.

So, experimental methods there are two methods. So, you can use the X-ray crystallography or the NMR spectroscopy and these are the two methods which you can use to determine the three dimensional structure of the proteins. For the X-ray crystallography, you can actually be able to have the very different steps. So, what you are going to do is first you are actually going to isolate the protein what which you are actually identity you want to identify the secondary or the tertiary structures. Then what you are going to do is you are going to purify this protein at 100% purity or more than 90% purity and once you have done the purify purification then what you are going to do is you are actually going to crystallize this protein. So, you are going to produce the crystals, and once you produce the crystal, then you are actually going to put these crystals for the diffraction and what you are going to get So, you are going to do the diffraction of these crystals and once you do the diffraction it is going to give you the diffraction patterns.

So, what is diffraction pattern diffraction pattern is actually going to his spots around the axis so it is going to give you the wherever the diffracted x-ray beam has hit the film and that is how it is actually going to give you a diffraction pattern. And ideally when you are want to collect the complete diffraction pattern of a protein it has to be rotate this crystal has to be rotated for an angle of 360 degrees because you can imagine this is the crystal you can have and this crystal actually has to be rotated for 360 degree then only you can be able to collect the diffraction pattern of or diffraction of all the electrons what

is present. Once you have the diffraction pattern or the diffraction data from the x-ray then you are actually going to put dust and you are going to analyze the diffraction data and that is how you are actually going to get the electron density map. Once you collect the electron density map it is actually going to give you the position of the electrons within the three dimensional. So once you got the electron density map then you are actually going to fit the protein molecules so you can do fit the protein molecules proteins molecule which means the protein sequence you are going to fit the protein sequence and once you are done the fitting then it is actually going to give you the 3D structure of the protein.

Once you got the 3D structure of the protein then you can actually do the quality assessment right you can do the quality assessment with the help of the three programs you can do the Ramchandran plot you can do the Pro-check and you can also do the Erato plot. So if you do all these kind of thing it is actually going to tell you whether the 3D structure what you have solved by fitting the protein sequence into the electron density map is correct or not. If you get the R factor which is called as the error factor which is approximately around 20 then you are going to say that the protein what you have solved or the protein structure what you have solved using the X-ray crystallography is very good. If you want to read more about since this course is not about the X-ray crystallography you will find a very good that kind of course on in the books or as well as you can actually get a lot of good resources if you are interested to understand each and every detail about the X-ray crystallography you can actually be able to even go through with this particular article and that actually will give you the very good idea about the X-ray crystallography as well as the NMR spectroscopy. Same to this NMR spectroscopy also has the different types of steps where you are first going to use the protein you are going to first the first step is you are going to purify the protein and you know that the normal protein what we are going to get from the bacteria is not going to be labeled not going to be NMR sensitive because it is so first thing is you are actually going to label the protein with the NMR sensitive nuclei's.

For example, you can use the N14 nitrogen you can use the carbon 14 carbon 14 and you can also use the hydrogen like the deuterium and so on and because of that the purified protein what you are going to produce is going to be NMR sensitive and then you are actually going to collect the NMR data and once you are going to collect the NMR data and you are going to analyze that data it is actually going to give you the protein structures and that protein structure is actually going to be called as the average structure. So this protein structure is called as average structure because the NMR is actually going to be performed in the liquid. So you can imagine that if I have a protein into the test tube, so this protein actually is freely moving and its domain and all other kind of structures are also moving because of that it is actually going to give me the

average structures. Apart from that you can also have the non experimental methods such as you can also do the homology modeling. So this is a useful and a fast structural solution method where the sequence similarity between the template and the target enzyme is used to model the 3D structure of the target enzyme.

The homology modeling exploits the idea that the amino acid sequence of a protein direct the folding of a molecule to adopt the suitable three dimensional conformation with the minimum energy. So what you are going to do is you are going to take so in this homology modeling, homology modeling depends on the reliability of or depend on the phenomena that the two proteins when they are actually having the similar kind of amino acids, they are actually going to adopt a similar type of folds because you know that the primary structure is actually going to direct the folding of these amino acids and that is why the when the primary structures are when the primary structure or the amino acid sequence is identical, it is actually going to fold into the same shape, which means if you want to use the homology model you are actually going to have the two things you are going to have a template structure. So you are going to have a template structure and you are also going to have the test amino acid sequence. Now first what you have to do is you are actually going to use this particular template and you are going to use this sequence and you are going to do the multiple sequence alignment. So you are going to do the sequence alignment and this sequence alignment is actually going to tell you whether this particular template is good for modeling this particular amino acid sequence or not.

Once this is done, then you are actually going to do two things, we are going to take the structure information from the template and on this structure information you are actually going to put the amino acid sequence what you have from the test and that is how you are actually going to prepare the model protein. So once you got the model protein, then what you are going to do is you are going to test the quality of the model. And we have already discussed the quality of the test of the model can be done by Ramachandran plot or you can use the ratah plot or you can do the pro check or you can do the verify 3D. These are the three or four different types of groups or the different types of programs what you can use and you can be able to do the error measurements. Once you have found that the structure is good, then what you can do is you can actually be able to utilize this model for the different types of applications.

The programs what you can use very oftenly is called as the modeler. You can use the modeler 9th version and that is actually going to allow you to do the modeling. So all these steps you can actually be used with the help of the modeler and then you can use the different types of programs to do the quality of the model verification with the Ramachandran plot and all these kind of things. So now let us move on to the methods to

determine the quaternary structures. So method to determine the quaternary structures. You know that the condition for the quaternary structure is that the protein should have the multiple subunits.

So if the protein has the multiple subunits then only you can be able to perform the quaternary structures. So how we can actually be able to know that the protein has the multiple subunits? So what you can do is you can calculate the protein molecule weight under the 2 conditions. One you can actually be able to calculate the protein's molecule weight under the negative conditions or you can actually be able to calculate the protein molecule weight under the denaturating conditions. So, when you denature in conditions, imagine that the molecular weight under the native condition is mol n. So, it is n right and under the denature in condition if the molecular weight is mol d then you can actually be able to calculate the oligomeric status by the molecular weight native divided by molecular weight denature right.

For example, so like molecule n versus molecule d. Let us take an example. For example, if I have calculated the molecular weight and if I calculated the native molecular weight is 120 kDa and if I have calculated the denatured molecular weight which is the 30 kDa then the oligomeric status would be the 120 divided by 30 which means the 4 which means it is a tetramer. So, once I calculated that it is actually a tetramer then I can be able to assure that there is a quaternary structure what is present. Now the question comes how you can be able to calculate the molecular weight of a protein under the native or the denaturing conditions. So calculation of molecular weight right. So for the native molecular weight what you can do is you can run the protein under the gel filtration chromatography.

So if you do the gel filtration chromatography although the scope does not allow you to explain the gel filtration chromatography but gel filtration chromatography is a chromatography technique which is actually going to filter the molecule based on the size. So if it is based on the size so if it is going to be 4 times because you are taking the 1 monomer and you are making the 4 monomers the size is going to be 4 times right. So that is why it is actually going to give you a pattern right it is going to give you a peak which is actually going to tell you that okay this is the size at which the protein is eluted. So this is called as the elution volume and utilizing this information and as well as the distribution coefficient you can be able to calculate the molecular weight of this particular protein and that is going to be called as native molecular weight. Now under what how to calculate the denatured molecular weight you can be able to calculate the denatured molecular weight with the help of the SDS page.

I am sure we all know about the SDS page so even if you run the protein on the SDS

page it is actually going to give you the molecular weight right. So you run first you run the molecular weight marker right so you run the marker and then you are going to run your proteins. So depending on the and then you calculate the Rf values for the these marker proteins right. So if for the individual marker protein when you calculate you are actually going to get the Rf values for each and every spot and then you can also be able to calculate the Rf value for the your protein and depending on the Rf value of this particular spot you can be able to calculate the denatured molecular weight. I have already discussed this whole thing in a and so that is how you can be able to calculate the denatured molecular weight and once you have the native molecular weight and you can have the denatured molecular weight you can be able to utilize that for calculating the oligomeric status which is the molecular weight native versus molecular weight denature right.

So with this we have discussed about the protein structures what we have discussed we have discussed about the different types of organization what is there in the protein structures. We have discussed about the the primary structures, we have discussed about the secondary structures, we have discussed about the tertiary structures, and we have discussed about the quaternary structures. We what we have so with this we have discussed about the protein structures and what we have discussed so far we have discussed about the primary structures, secondary structure, tertiary structure and quaternary structures. While we were discussing about the primary structures, we have also discussed about the method to determine the primary structures.

So, we have discussed about the Sanger's method or the Edman degradation methods. So, both the methods are utilizing the similar approach where you are actually going to label the terminal amino acids and then you are going to use the acid hydrolysis so that the terminal amino acid is going to be released and that terminal amino acid the labeled terminal acid going to be identified by running it onto the thin layer chromatography. And apart from that, we have also discussed about the secondary as well as the tertiary structures we have discussed about the methods to determine the secondary as well as the tertiary structures. And lastly, we have also discussed about the homology modeling we discussed about the X-ray as well as the NMR spectroscopy. And lastly, we have also discussed about how you can be able to determine whether the protein is going to have the quaternary structures or not.

So, with this, I would like to conclude my lecture here. In our subsequent lecture, we are going to discuss some more aspects related to protein. So, with this, I would like to conclude my lecture here. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 03
Basics of Biomolecules
Lecture-16 Enzymes

Hello everyone. This is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering, IIT, Guwahati. And what we were discussing, we were discussing about the properties of the different biomolecules which have extensive role in running the cellular metabolisms and they are also been responsible for different types of activities what are being done in the cells or in an organism actually. And majority of these biomolecules are be a part of the different types of activities and it is important for us to understand these structure and the function of these biomolecules so that you can be able to understand the molecular processes what they are actually going to run and regulate. So in this series, so far what we have discussed, we have discussed about the genetic material in the DNA.

Although we are going to discuss in detail about how people have discovered the DNA as the major genetic material what is present in the different types of organisms. But for the time being, you can imagine that the DNA is the major genetic material what is present in the different types of organism. There are exception that there are other kind of biomolecules which is also been responsible for carrying the genetic information from one generation to another generation. Apart from that DNA is also actively participate into replications and that is how the DNA is going to make multiple copies and these multiple copies are going to be given to the daughter cells after the cell division and the mitosis and meiosis.

And then apart from that the DNA is also going to be active in terms of the into another activity which is called as transcription and in the process of transcription the DNA is going to make the RNA and then these RNA are actually going to be responsible for making the proteins. So then subsequent to that we have also discussed about the RNA and the RNA is responsible for production of the proteins and the proteins are responsible for you know the different types of activities. So one of the major thing what we have discussed we have discussed about that the protein is the building block and if you recall in the previous lecture we discussed about the relevance of the proteins. So as a so we have in the previous lecture we discussed about structure and the function of the different protein molecules. So we discussed about the amino acids what are the different types of amino acids are present and then we also discuss about that the proteins is having the four layer of structures.

So it has a primary structure and secondary structures and tertiary structure and quaternary structures and very extensively we discuss how you can be able to study these different types of structures. So you can actually be able to have the different types of reactions what you can perform to sequence the reactions sequence the protein molecules and so on. Now in today's lecture we are going to discuss about the enzymes because enzyme is also proteinaceous in nature in majority the enzymes are proteinaceous in nature and these enzymes are actively being participated into the two different types of activities. First they are actually been responsible for running the metabolic reactions. So if you recall we have discussed about that the enzymes are actually responsible for running the catabolic reactions and the anabolic reactions.

Apart from that the enzymes are also responsible for running the detoxification reaction so that they can be able to produce the urea and then urea is actually going to be secreted out from the body in the form of the acids and the urine also and apart from that the catabolic reactions are responsible for energy production and this energy is actually going to be utilized for different types of activities. It is going to be utilized for growth, reproduction and the synthesis and the synthesis of the new biomolecule. The synthesis part is actually a part of the anabolic reactions. So it is actually going to participate into the anabolic reactions which is nothing but the synthesis and when we were discussing about the anabolic reactions we discussed that the anabolic reactions are so we discuss about the biosynthesis of the different types of amino acids and other things. So in today's lecture we will going to discuss about what are the different types of enzymes and what is the basic property of enzyme and apart from these biogasity role of these enzyme in the biochemistry the enzymes are also having the extensive role in the case of molecular biology protocols or the molecular biology processes.

So that the processes we are actually going to discuss in detail but today we are just going to discuss the different types of enzyme which are actually having the role in the molecular cloning. So what is the enzyme? So the enzymes as the name suggests is also being called as the biological catalyst. Okay so these are the catalyst molecule which are present in all the living cells. They are also present in all the living cells. What is their job? So just like any other chemical catalyst their job is to convert the substrate into the product and the enzymes are mostly been made up of the protein.

There is an exception that enzymes are also been made up of the RNA molecule. These are these enzymes are called as the ribozymes. So there are enzyme which are made up of the RNA molecule and these enzymes are called as ribozymes but this the amount of these enzymes are very very small and they are also being required for a specific function. So but majority of the enzymes are made up of the protein molecules and the as

I said you know enzymes are called as the biological catalyst. So what is the role of the catalyst? The role of the catalyst is that it is actually going to increase the rate of the chemical reaction right.

So it is actually going to increase the rate of chemical reaction so that the more number of substrate molecule will get converted into the product but at the end of the reaction it is actually going to remain unchanged which means it is going to perform the chemical reaction but it itself is not going to be a part of the reaction. That is why at the end of the reaction the enzyme is going to remain intact and that is how it is actually going to keep running the reaction right. Now the first question come is how the enzymes are working right. So what the enzyme is doing for example if I am taking an example that you have a substrate A it is getting converted into substrate B right. So what it is doing is it is actually you know converting A to B for example but like more precisely so and the molecule A is actually having a phosphate group okay and it is reacting with the molecule C for example and molecule C does not have a phosphate group but it has the groups right.

So and then it is forming the B molecule and B molecule is nothing but the C phosphate. This means it is actually breaking a bond here right and it is transferring this particular group on to the C molecule and that is how it is generating a B molecule which is nothing but the C phosphate. So it is here in the B it is actually making a bond. So at one place it is breaking the bond in the other side it is making bond okay. So how it is going to happen? It is going to happen that you have the substrate A plus B and that A plus B is actually you need to so what is happening here is that you are supposed to break some of the bonds and then you are supposed to make the new bonds right.

So if you want to break the bond you are actually going to infuse the energy into the system right and that is what it is actually going to happen. So if A and B are actually reacting with each other they are supposed to cross a barrier and once they cross the barrier then there will be a change of groups between the A and B and that is how they are actually going to form the PQ. This barrier is actually going to be high when you do not have the enzyme okay. This barrier is going to be on a higher side that means if you do not do this reaction without in the absence of enzyme then you are supposed to supply more amount of energy and as a result it is actually going to be difficult for this reaction to proceed. Whereas in the presence of enzyme what will happen is that it is actually going to lower down the amount of energy what you require to catalyze this reaction and this is what exactly what is going to happen when you have an enzyme.

So enzyme as a catalyst change the rate of chemical reaction but it does not alter the equilibrium okay. This means it is actually going to reduce the activation energy which

means the energy what is required to activate to A plus B to form the P plus Q is actually going to be on a lower side and that is how it is actually going to achieve this value much quicker right and that is how it is actually going to catalyze more number of reaction. So a catalyst function by lowering the activation energy of a reaction the energy barrier for a reactant to become the product. Now this is what exactly what happens right. Now the question comes why we need the enzymes.

So you need an enzyme because of this reason right. So you see there are some examples of the different types of enzyme and I am giving you the value of the enzyme constant right the rate of reactions in the presence and absence of the enzyme. So this is the absence of the enzyme and this is the presence of enzyme. So what you see here is that in the absence of enzyme if you are going to perform the reaction what is being catalyzed by the carbonic anhydrides it is actually going to have a reaction rate of 1.3 into 10 to power minus 1 okay.

Whereas if you are adding the enzyme to this reaction it is actually going to have the rate of reaction as 1 into 10 to power 6 which means you are actually going to have the rate enhancement if you have the enzyme as in the range of 10 to power 6. So 10 to power 6 fold there is a rate enhancement same is true for the course mate mutase if you have the non enzymatic reaction it is going to be in the range of 10 to power minus 5. But if it is a enzyme catalyzed reaction then the rate of reaction is 50 this means there will be an enhancement of 10 to power 6 folds. Same is true for the triosulfate isobanase in the absence of enzyme it is going to have a rate constant of 10 to power minus 6 per minute whereas in the pace of enzyme it is going to be 4300. So there will be an enhancement of 10 to power 9 folds.

So what it says is that enzyme which is going to increase the rate of reactions but it will not participate into the reactions and it will going to be remain unutilized at the end of the reactions. So there are other example also right. So why we need an enzyme? We need an enzyme to increase the rate of reactions increase the rate of reaction that is one point. The second point is that some of these reactions will not going to proceed at a rapid rate. So to perform these reaction at a rapid rate you are also going to increase the temperature.

So for example if you are to run the carbonic and HEDIS reactions right you are supposed to have a degree angle if you supposed to have the heating of these reactions at 100 degree Celsius and couple of atmospheric pressures. Those kind of conditions are non physiological conditions. So if you are going to perform the same reactions you are actually going to have the non physiological conditions like for example 100 degree Celsius. Most of the organisms will not going to survive at 100 degree Celsius right. So

what is the permissive temperature? Permissive temperature is 37 degree Celsius.

So if I supposed to run the reaction at 37 degree Celsius I am supposed to enhance the rate of reactions and that rate enhancement of the rate of reaction is only possible if I have an enzyme into the reactions right. Now what are the different properties of enzyme? So enzymes are made up of the protein except there is an exception of ribosome. So that we have already discussed that it is made up of the RNA. Then due to presence of the amino acids it provides the specific environment for catalyzing the reaction with the different types of substance. You know that the since the enzymes are made up of the proteins and the proteins are made up of the amino acid and you have the 20 different types of amino acids and all these amino acids vary in terms of the different types of property like the charge, polarity, then you also have the hydrophobic amino acids, you have hydrophilic amino acids and so on.

And all these actually provides a local micro environment which is different for the different types of substrate and that's why the enzyme can be able to recognize even the subtle changes into the substrate. That's why they are very precise and they are very specific. The substrate binds to a small pocket within an enzyme. This pocket is known as the active site. The molecule produced by the reaction is called the product right and the enzyme catalyzed reactions are very rapid than the uncatalyzed reactions.

Now they are very specific towards the substrate and the product and the enzyme activity can be modulated by the non substrate molecules such as allosteric control or the covalent enzyme modifications. In a few specific cases, enzyme amount can be modulated by the synthesis or the degradation. So this actually happens within the cell when they are actually either degrading the enzyme so that they can be able to modulate the reactions or they can be able to synthesize the new molecule. So that anyway we are going to discuss when we are going to discuss about the translation and other kinds of molecular events. Now how the enzyme is recognizing the substrate? So it is actually recognizing the substrate due to three important parameters.

One is geometrical complementarity, the second is electronic complementarity and the third is stereospecificity. Geometric complementarity means it is actually going to see whether the 3D structure of the substrate is matching with the enzyme or not right. So you can see that this is a substrate and it is matching exactly with the 3D structure of the enzymes or 3D structure of the active site. The third is electronic complementarity. So electronic complementarity means whether the electron donor and electron acceptor groups are being compatible to each other which means wherever the electron donor is present onto the substrate whether the electron acceptor is present on the enzyme or not because you have two pairs, one is enzyme, the second is the substrate.

So if you have the electron donor onto the substrate molecule then you should have the electron acceptor onto the enzyme actually. For example here right this is the hydrogen donor right so you have the hydrogen acceptor on this. So when the enzyme will when the substrate will fit into this cavity the hydrogen donor and the hydrogen acceptor they will actually going to interact with each other and that is how there will be hydrogen bonding formations. Similarly you have the one charge right so you have the for example you have the negative charge. So it is actually going to interact with the positive charge what is present onto the enzyme.

So there will be a salvage interaction between the substrate and product. That is how it is actually going to bind very strongly to the enzyme. And there are other kinds of interaction also for example you have the hydrophobic interaction, hydrophobic molecule, hydrophobic substances what is present onto the substrate and then you also going to have the hydrophobic groups onto present on the enzyme. So these are the some of the things what is responsible for the substrate specificity. Apart from that you also require the stereosensitivity which means that it will actually going to recognize whether the substrate is L type or the D type.

Although this particular in this particular course we are not going to discuss in detail about any of these aspects because we are actually going to discuss more about the molecular biology related stuff. So if you want to know more about these things you can we have another MOOCs course which is called as enzyme science and technology and you can actually be able to follow that. So there is a MOOCs course where you can actually be able to use that. So this is what exactly it says that the geometric complementarity means that enzyme binding side has a structure which is complementary to the substrate it binds. Then electronic complementarity the amino acid that are from the that is forming the enzyme binding side are arranged to specifically interact and attract the substrate molecule and then the stereosensitivity that the binding of chiral substrate and the catalysis of the action is highly specific due to the large part of inherent chirality of the L amino acid that comprising enzyme.

So if I summarize the properties of the enzyme the enzymes are actually going to have the different types of groups what are present and enzymes are actually going to be very specific for their substrate. Apart from that the enzymes are also requiring the metallic and as well as the other kinds of small groups which are actually being a part of the cofactor. So if it is a metal then it is going to be called as cofactor and if it is a small molecules then it is going to be called as coenzyme. So these are just some examples of the cofactors for example we have the copper, iron, potassium, magnesium, manganese, nickel, selenium, zinc and for example the copper is a cofactor in the cytochrome C

cytochrome C oxidase. Then for iron it is actually present as a cofactor in catalase and peroxidase, potassium is present in pyruvate kinase, magnesium is present in hexokinase, glucose 6 phosphate, manganese is present in the arginase, diphosphorylase and urease, nickel is present in urease and so on.

Similarly we have the co-enzymes. So coenzymes are mostly the vitamins or the other kinds of molecules. So it is a small organic molecule for example you have the biocytidine so that will be a coenzyme for the carbon dioxide right. So it is actually going to bind the carbon dioxide then we have coenzyme A, it is actually going to have the acyl group. Then we have coenzyme B12 so it is going to have the hydrogen and alkyl groups. Then we have FAD which is going to have the electron and so on.

So these are the some of the things and since the enzyme require these molecules for their optimal activity if these molecules are not present then they will be responsible for different types of disease. For example if you have the deficiency of the iron then you are going to have the parenthesis anemia and so on. And then pellagra is being caused by some of the vitamins deficiency and so on. Now we will focus more on the enzyme which is responsible or which are going to participate actively into the molecular cloning. So these are the general you know the scheme of the molecular cloning where from the genome you are actually going to identify the gene and you are going to amplify this gene with the help of the polymerase.

This process is called as the polymerase chain reaction and these are the things we are actually going to discuss in this particular course. So just for you know this is summary of what we are going to discuss in molecular cloning. Then you are going to digest this with the restriction enzymes and that is actually going to generate the cohesive ends in both the sides. The same is true for the plasmid also and then you are going to have the cohesive ends of the plasmid and then you are actually going to do the ligation reactions. And once you do the ligation reaction you are going to have the ligated plasmids.

So these are the recombinant plasmids and then you are going to transform this and that is how you are actually going to have the you know the organisms the transformer organism and that can be used for protein production. So this is the general scheme and in this particular scheme what you see here is that you are first using the polymerase. So this is the enzyme 1 what you require. Then you also require the restriction enzyme. So this is the enzyme 2 and then you also require the ligation reaction.

So you also require a third enzyme which is called as ligase. So these are the three enzyme which are very very crucial for the different types of but these are the some of the enzyme which is actively participating into the different types of activities within the

molecular biology. So the first is polymerase right which is required for the PCR amplifications. Then you require the repetition enzyme that is for the cutting the DNA at a specific site. Then you also require the alkaline phosphate that is required for the removal of terminal phosphate group and then you also require the DNA ligase which is joining of the two DNA strands.

So we will start first with the restriction enzyme then we are going to discuss about the polymerase chain reaction polymerases because that we are going to cover when we are going to discuss about the PCR and then we are going to discuss about the alkaline phosphatase and at the end we are going to discuss about DNA polymerase DNA like research. So the first is restriction methylase system. Substitution methylase system is immune system which is present in the prokaryotic system ok and it does not allow the propagation of foreign DNA ok. So this it distinguishes with self versus foreign DNA. So foreign DNA is the DNA from the infectious organisms, self will be that the DNA from your own right.

So although the precise mechanism or distinction is not known but based on the available literature in the absence of methylation a closed complex is formed and allow the proper activation of the cleavage activity of the enzyme. The presence of methyl group on nucleotide does not allow the formation of the closed complex and consequently the enzyme falls from the DNA. So this is what exactly happens. So restriction enzyme or restriction methylase system what it is doing is that it is checking the DNA for the presence of the methylation on to the adenine groups.

So you have the methyl group right. So methyl group either would be unmethylated, hemi methylated or the fully methylated. So for example in this case this is the one DNA is unmethylated the other DNA is methylated right. So this is unmethylated DNA. Now if I add the restriction enzyme what will happen is that it is actually going to recognize that okay there is no methylation and that is how it is actually going to cleave. So it is actually going to cleave this band more and that DNA and that is how it is actually going to generate the sticky ends okay.

But if there is a methylation for example in this case right this is the fully methylated then the enzyme will not be able to bind and as a result there will be no degradation of the enzyme. So this DNA is going to be considered as self-DNA and this is actually going to be considered as non-self DNA. And this is a kind of a defense response what is present in the lower into the prokaryotic system but this system has a very unique feature that it actually generates the sticky ends. So you have different types of restriction enzymes and they actually generate some of these restriction enzymes are actually generating the sticky ends and that can be used into the molecular cloning. Now

the question comes how the restriction methylase system actually recognizes the cleavage site.

So for this for example we have the four restriction sites, four restriction site RE1, RE2, RE3, RE4 and now what we have done is we have added one restriction enzyme right. So what will happen is that the restriction enzyme will actually go and non-specifically will go and bind the DNA right at multiple places okay. So it will go and bind a non-specific binding to the restriction sites. So it is going to bind to site 1, site 2, site 3, site 4 and every sublase what it actually going to do is it is actually going to look for the binding cause or binding right. So it can have two different types of binding either it can have the loose binding or it can actually have the tight binding.

If it is having the loose binding then the DNA will not be able to cleave. So it is naturally going to no cut okay. If it is a tight binding or then it is actually going to cut the DNA okay. So what happened is that if there is a closed complex or the tight binding then the enzyme will sit and then it is actually going to catalyze the cleavage reactions and as a result it is actually going to generate the degraded DNA and then the enzyme is actually going to be released. However, the estrogen enzymes are also an enzyme made up of a protein.

So they are not going to utilize into the reactions and ultimately it is actually going to come out from the reactions. Now once the people have started discovering the different types of restriction enzyme they also put rules and regulation for putting the name of these enzymes. So the nomenclature of a restriction enzyme due to the extensive search of the presence of restriction enzyme in the different molecular organism a nomenclature system has been adopted. In this system the first alphabet represent the name of the genus, the second alphabet represent the species, third alphabet gives the information about the strain and the fourth is the order in which the enzyme has been isolated from the particular microorganism.

For example, this is an enzyme name EcoR1. So this is the restriction enzyme. So here the first alphabet is E, the second alphabet is CO, the third is R and the fourth is 1. So for the E it stands for Escherichia.

So this means EcoR1 is being discovered from the E. coli. Then CO which is going to be the species, so coli species. And then R is for the strain. So it is actually been isolated from a strain called RY13. So this is the strain and the one is that it is the first restriction enzyme from there which has been discovered from this particular enzyme.

Now so different types of restriction enzymes. So restriction enzyme vary in restriction

cutting site and cofactor requirement. So you can have the type 1 restriction enzyme, you can have type 2 restriction enzyme and you can have the type 3 restriction enzyme. So for the type 1 restriction enzyme, the restriction site of type 1 restriction enzyme consists of 3 to 4 nucleotides at 3 prime and followed by a non-specific stretch of 6 to 8 nucleotides and a 4 nucleotides at 4 prime. And for type 3 restriction enzyme, the restriction site has two separate non-pilomatomic sequence arranged, inversely arranged and the cutting site is 20 to 30 base pair away from restriction site. So the type 2 restriction enzyme is composed of two subunit, RES and MOD.

The MOD subunit is required for the modification whereas the RES is required for the cutting the unmethylated DNA. Then we have the type 2 restriction enzyme. So the type 2 restriction enzymes are very useful for monotonic cloning because they generate the sticky ends. So the recognition site of type 2 restriction enzyme is 4 to 8 nucleotide long and it cuts the DNA within the specific site. Due to this feature, the type 2 restriction enzymes have an application in genetic engineering for cloning purpose.

It is composed of three different types of subunits, M subunit, R subunit and S subunit where the M is required for methylation, R is required for cutting the DNA and S is recognizing the sequence which it actually going to be specific. And type 2 restriction enzymes are further being classified into type 2E, type 2B and all that that we are not going to discuss in detail. So this is just a summary of the different types of restriction enzymes. So we have the type 1 restriction enzyme, type 2 restriction enzyme and type 3 restriction enzyme and this is just a summary of different properties which you are actually going to use. Now we are going to have the properties of the restriction enzyme because these are the restriction enzyme which are actually going to be having an extensive role in molecular cloning.

So first is they are actually going to have the palindromic sequence. The restriction sequence of type 2 restriction enzyme is palindromic in nature. It means that the sequence read out would be same in forward and reverse direction. For example, the Bamach one had an restriction site which is GGATTC. So let us see what is mean by the palindromic sequence.

So GGATTC, so this is 5 prime end, this is the 3 prime end. Now if I write the reverse sequence, what is the reverse sequence? CCTAAG. So now we see that if I read this in this direction say GGA, if I read this sequence in this, it is going to be say that GGA. So GGA, GGA, TTC, TTC. So if I read this sequence either in this direction or either in this direction, it is actually going to be same and that is why these kind of sequences are going to be called as palindromic sequence. Now what is the advantage of this? That advantage is that it is actually going to be recognized by the restriction enzyme.

The second is it is actually going to generate the sticky end. The type 2 restriction enzyme cut both DNA strand together to generate the DNA with the hanging DNA stretch with the 4 to 8 nucleotide. This DNA stretch contain fragments are cohesive to each other as sequence present on complex 1 will be complementary to the sequence present on the complex 2. So for example, this is the DNA. So this is the DNA.

Now this is 5 prime, this is 3 prime, this is 3 prime, this is 5 prime. And if this is the restriction, this is the recognition sequence. Now if I cut this DNA, what I will generate is this. What I will generate into this DNA, with this DNA this portion will go. And whereas with this DNA, it is actually going to go with this.

So what you see here is this is actually sticky to each other. This is going to be sticky to each other. This is what it is actually going to happen. This is the 5 prime, this is 3 prime, and this is 3 prime, this is 5 prime. So if I put this together again, they will actually go and stick like this and then this small gap will be filled and that gap is actually going to be filled by the ligase enzymes.

If I want to set up the restriction reactions, then I can use this. So I can have the DNA, restriction enzymes, buffer, BSA and esterile water. And the restriction reactions are supposed to be put in a larger volume so that you should have the free access of the enzyme. So you can do the in esterile water. So in esterile water, what you are going to do is first you are going to add the buffer. So first you are going to take the water, then second you are going to take the buffer, then you are going to add the enzyme.

Okay. And that and then you are actually going to make the buffer. So that is actually going to have the restriction enzyme master mix. Now to this master mix, if I suppose have 5 different types of DNA, right, so I can add the DNA, right. So I can add the vector, I can add the template and all that. And that's how I can set up the different reactions.

And then I incubate that on 37 degrees Celsius for 12 to 18 hours, right. And at the end, it is going to cleave all the, from the restriction sites and that's how I am going to get the cohesivants. And then I can just put the vector and the fragments together and then it is going to be done with the ligation reactions. Now the second enzyme what we are going to discuss is the ligase. So ligase is joining the two DNA fragments to generate the chimeric DNA is the basis of the cloning.

It is an essential step to generate the clone containing foreign DNA in a vector. When the cohesivant generate by the action of the precision endonuclease on a DNA associated

with each other, a nick remain to seal and give complete circular DNA. So what DNA ligase is doing? It is an enzyme which requires the ATP or the NAD plus as a cofactor to catalyze the ligation reactions. Ligase is processing ATP to generate AMP and then AMP is making an adduct with the enzyme to form the ligase AMP complex. This complex is binding to the 3 prime and 5 prime of the DNA bearing nicks and bringing them together. AMP is released and phosphodiester linkage is formed between the 3 prime and 5 prime end to seal the nick.

So this is what exactly it is going to do, right. So when you have the enzyme, right, you have the T4 DNA ligase which utilizes the ATP or you can have E. coli DNA ligase which utilizes the NAD plus as a cofactor. And either of these cases, suppose this is the DNA and you have a nick here, right. So you have a nick here, right. And then what will happen is that the enzyme is actually going to make a complex with the ligand or like with the cofactor.

So it is going to form the enzyme AMP complex and then enzyme is actually going to bind this particular sequence and it is actually going to supply the AMP, right. And then AMP is actually going to come off and then there will be a bond which is going to be formed between the phosphate and the OH which is present. And that is how it is actually going to seal this particular nick. Now how we can be able to set up the ligation reactions? So for ligation reaction what you require, you require the vector or DNA 1, then you require the insert which is DNA 2, then you require the ligase buffer, then you require the BSA and then you also require this trial water.

Remember that the reticence enzyme reactions are supposed to be done in larger volumes. That is why we have set up that reaction in 50 microliter whereas the ligation reaction supposed to be done in a smaller volume so that there will be higher probability of these fragments interacting with each other and making the ligations. So in a ligation reaction what you are going to do is you are going to take the trial water first. So you are going to take the trial water first, then you are going to add the ligase buffer, then you are going to add the BSA and then you are going to add the ligase, T4 DNA ligase, 5 to 10 units per reactions and then you are going to add the DNA. If it is a vector and insert then it has to be added in a 1 is to 3 ratio so that there will be higher probability that the vector will interact with the insert and it is going to form the chimeric DNA. Once you set up the reaction, then these reactions have to be set up incubated at lower temperature, 16 degrees Celsius so that it is actually going to help in forming the hydrogen bonding and that helps in sealing the leaks.

And then you incubate that on 60 to 24 hours after that you are going to transform this ligation reactions into the suitable host and that is how you are going to get the

recombinant DNA. The third enzyme what we are going to talk about is the alkaline phosphatase. So alkaline phosphatase is required when you want to do the directional cloning. So the digested linear plasmid containing the cohesive end on both the side with the phosphate has a tendency to recirculate which means if you have a vector, if you have a vector and suppose you have the Eco R1 on this side and you also have an Eco R1 with this side and suppose you have insert that also being digested with Eco R1, then you have a two probability either this vector will go, this insert will go and sit here or this vector itself is actually going to surprise with each other.

So you are actually going to have the two possibilities. One that it is actually going to take up the insert and it is going to form the recombinant DNA. The second that the vector itself is actually going to be get sealed with each other and if that happens, then it is actually going to be recircularization and it is not going to give you the recombinant DNA whereas if the insert comes here, then you are going to get the recombinant clone. So removing the terminal phosphate group prevents this possibility and for this purpose, the alkaline phosphatase is used. So if I want to avoid this, what I can do is I can just remove this particular phosphate group because on this side it has a phosphate group, on this side it has OH group. So if I remove the phosphate group, it is actually going to have the OH on this side, OH on this side and if that happens, then this particular fragment will not be able to be surprise on its own.

It actually requires the supply of this phosphate group and it also requires the help of the ligase reactions. So alkaline phosphatase removes the 5 prime terminal phosphate groups and in that condition only in the presence of insert DNA as its supply phosphate group HN to provide the ligation reactions. So this is what exactly I was trying to explain. You have the vector which has a phosphate group and you have OH group on this side and then you have OH group on this side and phosphate group on this side.

So if you put the ligation reaction, it is actually going to circularize. So you are going to get the plasmid back. So instead of, but you do not want the plasmid back. You want a recombinant clone. So in that case what you do is you treat this with alkaline phosphatase. So if you treat the alkaline phosphatase, it is actually going to chew up all the phosphate what is present on the termini.

So as a result, it is actually going to have the OH on the other side, OH on all the sides. Now this cannot recircularize even if you put it onto the ligation reactions. So there will be no ligation. Now if I put this along with the insert, insert has the phosphate group which is present. So insert will sit here and then it is actually going to have the nicks and these nicks are actually going to be sealed by the ligase and that is how it is actually going to give you the ligation reaction and these ligated product can be transformed into

the bacteria or the other force and that will going to give you the recombinant or the recombinant flows.

So these are the some of the enzymes what we have discussed. So we have discussed about the enzymes what is in general. So what are the different properties of an enzyme and how the enzyme actually works in the biological system and then at the end we have also discussed about the different types of enzymes and their properties which are actively participating into the different types of molecular cloning reactions. So in this particular module what we have discussed, we had discussed about the structure and function of the different biomolecules which are actively participating and regulating the different types of biological pathways and different types of biological properties or biological actually pathways. And so with this I would like to conclude my lecture here.

In our subsequent lecture we are going to discuss some more aspects of related to molecular biology. Thank you. 1

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 04
Central Dogma of Molecular Biology
Lecture-17 Genetic Material (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from the Department of Biosciences and Bioengineering IIT Guwahati. And what we were discussing we were discussing about the different properties of the cell in the course molecular biology. So far what we have discussed we have discussed about the basic properties of the biological system where we have discussed about the cellular structures, we discussed how the cells are dividing and then we also discussed how the cells are dying through a process known as apoptosis. And in the previous module we were discussing about the different types of bio molecules. So we have discussed about the DNA, we have discussed about RNA and we also discussed about the protein and the enzyme.

So with this brief discussion about the biological system we have discussed about the bio molecules we would like to you know ask the questions how the information from the one generation to another generation is passing and what are the different types of molecules which could be responsible for passing the information from the one generation to another generation. Now what you see is that I am sure you might have noticed that some of your own traits are matching with your parents. Similarly the traits what are present in the plants they are also of mixed traits right they are also having the some information from the one parent and the some information from the other parents. And on the other hand you might have seen that some of the diseases which are which are propagating within a particular type of families.

For example, I am sure you might have noticed the traits like height, eye color and other kinds of phenomena like the hair colors and the way you speak and the way you actually you know behave is all being transferred from the parents to the child. Similarly you might have seen that the seeds from a red color flowering plant always produce red color flowers naturally right and then you might have seen also the different kinds of variation like if you have the cross breeding of a white flower and the red flower you will see that they were actually having the pink flower and so on. And there is a classical examples of the hemophilia also where which actually runs in different types of families from generation to generation. Now the first question comes that how does it happens and the answer to this question is that it is all because of the hereditary. So what is heredity? So heredity or also called as inheritance is the passing on the traits from the parental generation to the offsprings either through the asexual reproduction or the sexual

reproduction.

So the offspring of the cells obtain the genetic information of their parents. So because of the hereditary or the inheritance you are acquiring the traits or you are acquiring the phenomena or acquiring the phenotypes from your parents. Definitely when it is if it is asexual reproduction the traits are going to be completely 100% intact. But if it is asexual reproduction then it is actually going to be mixed because you are going to have the 50% traits from your mother and the 50% traits from your father actually. And depending upon who which trait is dominating and which type is recessive it is actually going to show you the phenotype.

So we are not going to get into the detail of the hereditary or we are not going to get into the genetics part. So because that is beyond the scope of this particular course and that be that for that you can actually go through some of the most courses on the genetic material on the genetics itself and that may help you to understand this particular phenomena that how the traits are getting you know expressed in some generation and they are actually being not expressed in other generation as well. So the first question comes that if it is actually the thing that you are actually going to acquire the traits from the mother and the father how this particular type of phenomena is happening and who is responsible for that. So who is responsible for carrying the information from one generation to another generation and the responsible the molecule which is responsible for this particular information or what for this particular phenomena is called as the genome or I will say genetic material because it is actually going from one generation to another generation and because of that this particular molecule is being called as the genetic material not the genome actually. And I am sure when we were discussing about the cell and even in the previous lecture also when we were talking about the biomolecule if you see the cell the cell is very complicated right it actually has a nucleus it has the cytoplasm it actually has different types of organ is in the prokaryotic system also you have all those things but not the membrane bound right.

So you also have the miniaturized level of or the primitive conditions of the electron transport chains you also have the other kinds of things. So basically if you think about the wherever the information can be stored there are multiple possibilities one is in a cell it can be nucleus or I will say the genome or I will say DNA because we know that the nucleus contains the genome which is made up of the DNA. The second possibility is the cytosol or I will say cytosol actually right and cytosol is mainly been contain one molecule which is called as protein and plus RNA right. So RNA is also present in cytosol then you also have the membrane bound organelles. So membrane bound organelles are also been made up of the protein and RN lipids right because the membrane and lipid right.

So we basically have the candidate molecules which are responsible for carrying the information but because the technique was not evolved or how the people have find out the who is actually responsible for carrying the information is a very long journey right. So let us first so what are the molecules we have? We have the DNA we have the potential target potential molecules like the DNA we have the proteins we have the RNA and these are the molecule which are actually having the sequence which having the stored information. For example you know that the DNA is made up of the nucleotides so that also is actually providing the sequence of nucleotides so that also can actually carry the in homogeneous amount of information. I am sure you can calculate if you have 4 nucleotides how many different types of random combinations could be possible to give you the different types of random DNA sequences. Similarly you have the protein which are made up of the 20 amino acids and all these 20 amino acids random combinations can give you the enormous information so that is how and the RNA also right the same way the RNA.

So these 3 molecules are actually having the similar kind of nature apart from that you also have the lipid but lipid does not have that kind of flexibility of storing the information. So that is why the lipid is straightforward is being discarded by the scientist what they were focusing on the DNA protein and RNA. But before that the people have done the crude experiment because until the people have not discovered the DNA protein and RNA they were under the you know the they were not having the technology. So they were doing the experiment they were trying to identify the carrier molecules which actually carries the information from one generation to another generation and that is how they have done multiple type of experiment. So let us see how we have you know we have done the what is the history of identifying the genetic material and then we will talk going to discuss about the different types of classical experiments how people have you know figure out whether the DNA is the genetic material or protein or RNA.

So the first experiment is been done by the Astrea scientist or I will say the priest Gert-John Wendell. So he is considered to be the father of genetics and he actually has done extensive experiment with the pea plants and where he has taken the combination of traits and that is how we have come up with the classical rules of the genetic the classical rules of the genetics and so what is the history of the genetic material. So it is well known that the qualities are passed down from one generation to the next the offspring shares certain characteristic with both of their parents. But the question is who is responsible for this? It was first time on the basis of its interest in the plant hybridization studies on the sweet pea the piscine satayabam monk in a monastery in Austria Gregor John Mendel try to find out this answer he proposed that the some factors. So in when the Gregor John Mendel was doing the experiment there was no technology of

involvement there were no technique available to say that DNA, RNA or protein but he said that there are factors and he used the term factors that carry the information on the manifestation of a characteristic or phenotype and how the traits are being passed from one generation to another generation.

So since he gave the clue about this particular factor and he said that this factor could be dominant factor and recessive factors and so on and all that and all these law of genetics you might have studied in some of the textbooks. So then people have started you know identifying this particular factor. So in the late 90s century by the three biologists, Thumodouvis, Karl Kornes and the Eric Vaughan worked on the Mendel works and proposed that different characters have individual hereditary carriers and the inheritance of specific traits in an organism called particles. So Dviris actually called these units as the pan genes and he actually come up with the theorem of the theory also that so he actually discovered that these factors which the Mendel was talking about is nothing but it is actually the pan genes which actually go and if the pan genes can move from one generation to another generation they are actually going to carry the that particular characteristic or the quality. Then in the almost 20 years later when the William Johnson and William Batson actually proposed the term gene and the genetics respectively but Edwards, Tranberger and the other continued to refer to the basic physical and functional unit of hereditary as the pan gene.

So basically from the pan gene it becomes gene and the gene was being considered to be a responsible factor. So genes are present in chromosome which are evenly distributed between the two daughter cells during the cell division and the biochemical study showed that the chromosomes are consist of protein and the DNA. So it is clear that the gene is present in chromosome and we are going to discuss about this nuclear packing and all that when we are going to talk about genetic material and the chromosome is made up of two parts the protein part or the DNA part. So the first question is that is the genetic material proteins or the DNA which means out of this chromosome which one is more responsible the protein part or the DNA part because it has the both the components. So until 1940s the proteins were thought of a genetic material because proteins are polymer made up of the 20 different types of amino acids which are abundant and encode the diverse information and you can easily calculate in fact you can actually go with that activity right.

How many different types of amino acid compositions or how many amino acid combination could be possible if you have the 20 different types of amino acids. So for example you have a very small protein of 100 amino acid if you have a small protein of 100 amino acid and you can have the 20 different random combinations you can actually be able to calculate that number that number is going to be very very big actually.

However based on the certain experiments that have been conducted from the time it was finally shown that the DNA not the protein actually carry the genetic material but we are actually going to discuss in detail about this particular these experiments and conclusion comes that it is a DNA which is actually being the molecule present within the chromosome and that is actually going to carry the information from one generation to another generation. Now the first question comes what could be the properties of the genetic material what could be the possible or the probable properties you should have or a molecule should have then only you can say okay this is the genetic material. So the properties of the genetic material okay so for a molecule to be considered as the genetic material because you can have the some requisite parameters to accomplish the its task right what is the task? Task is to carry the information from one generation to next generation right that is the task of a genetic material and that is why it should have a repeat circuit conditions.

What are the conditions number 1 is stability number 2 is it should be editable and expression number 3 it should have the mutations and a number 4 it should be getting replicated. So stability it must contain all the biologically useful information in a stable form which means it should be stable it should be you know it should be the distance for any kind of damages and if it there is a damage it should have a mechanism to recover from that particular damages then it should be having the irritability and the expression. So it should process a hereditary units which follow the Mendelian inheritance and control the expression of a particular phenotype. So it should have the components we should have the and I think all these you are going to understand when we are going to talk about the transcription and translation that how there are different components which are present within the genetic material and that actually controls the expression of a particular gene. And so you should have these kind of switches you can have these kind of switches so that you can actually be able to modulate the expression of the particular gene and that is how you are actually going to modulate the overall phenotype of that particular organisms.

I am sure you might have noticed when you are going into the sun and it is very hot outside you always start sweating. So that sweating occurs because there is a you know there is a expression profiling changes within your body and that is how they started you know throwing the sweat. Same is true when you are entering into a you know into a cold room or entering into a place where you have AC and then you actually you know all your set is disappeared and then you also feel cold right and when you feel cold you find that the skin is actually you know becoming more you know the contract actually. So these things actually been you know done simply by the information what is present inside the genetic material. Then the third property is about the mutations right.

So genetic material is also going to acquire the mutations some of the mutations could be good some of the mutations could be wrong and that is how the mutation accumulation of mutation is actually going to result into a change in the phenotype and that will be responsible for evolution right. So mutation is the random change which may occur and that may be a chance for evolution actually. And then number four is replication because as I said you know genetic material is you are going to have only one copy of genetic material. So that genetic material has to be replicated and that is why you are going to have two copies of genetic material and then you can actually be able to in a situation that you can be able to share. Remember that when we were talking about the cell cycle we said that during the S phase the DNA is actually going to be duplicated right and then only it is actually going to be shared between the cell.

And the same is true for between the parents and the offspring as well that you are going to have the two copy of genome and then when you are going to share one copy with your offspring and the so on. So that should have the ability to synthesize its own copies. This replication is actually what we are going to discuss in our subsequent module we are going to talk about the replication transcription and translation. Now there are some direct evidences or experimental evidences and there are some indirect evidences which prove that the DNA as a genetic material. So there are direct as well as indirect evidences which have ruled out the protein or the RNA as a genetic material and that has proved that the DNA is actually the most acceptable genetic material which is present starting from the prokaryotes to mammals.

There are exceptions and these exceptions also we are going to discuss. So in 1928, Friedrich Griffith actually performed an experiment for the bacterial transformation and by doing these experiments he proved that actually the information what is being present in the DNA is actually carried from one bacteria to another bacteria and that is how it is actually going to change its phenotype and that is how it is actually going to be responsible for the death of the mice. Then in 1944, the Oswald, Avery and McLeod and McCarty actually done the experiment on the transformations and that also has proved that it is actually going to be the DNA which is going to be responsible factor. And then we have the Alfred Hershey and the Chase experiment on the T-even bacteriophage and that is how it that also has proved that the DNA is a genetic material. So if you go by the timeline in 1869, the Friedrich-Mateur actually isolated the nucleic acid or I will say the genetic material from a white blood cells.

Then in 1928, so, almost after 50 years or 60 years, the Friedrich Griffith actually demonstrated that the genetic information from the one bacteria goes into another bacteria through a phenomena which is called as the transformations. Then in the year of 1944, the McLeod and McCarty actually identified that the DNA is actually a

transforming agents, it actually carries the genetic information and then it actually can change the phenotype of the other bacteria. Then in the case of 1952, Hershey and Chase actually confirmed that the DNA is a genetic material with the help of the viruses. And then in 1962, the Watson, Crick and all those people have got the Nobel Prize because when they have discovered the structure of the DNA. So, let us first discuss about the Griffith's experiment to understand that the DNA is the genetic material and then we will talk about the McLeod and McCarty's experiment and then ultimate and at the end we are going to discuss about the Hershey and Chase experiments.

So, the Griffith's experiment, so, he has used a bacteria which is called as pneumococcus. This pneumococcus is actually called so, he has used the bacteria which is called as streptococcus pneumonia and he has used the two different types of strain, S strain and the R strain. So, S strain and the R strain is also going to be called as, so, the S strain which is a virulent pathogenic strain that is S strain because called as smooth strain or the S third strain and the R strain is a recessive strain. So, R strain is the a virulent strain and it is non-pathogenic strain known as the R virulent or the rough strain or the R2. So, S strain is actually going to cause the disease whereas R is not going to cause a disease.

So, if the S is going to cause a disease S is actually going to kill the mice's whereas R is not going to kill the mice's. So, S strain has a smooth outer coat of the polysaccharides and the R strain lacks this polysaccharide coat and therefore, its surface appears rough. S3 strain was virulent possessed a lipopolysaccharide capsule and could kill mice by causing a disease pneumonia and made round colonies on a cultural plate. Whereas the R2 strain was a virulent and lacked a lipopolysaccharide capsule giving life to the rough shape colonies onto the cultural plate. So, these are the some of the properties given.

So, serotypically it can be SR2 or the S3 and morphologically the R strain is a rough strain whereas the S strain is a smooth strain. Then capsule was absent in the case of R strain whereas it is present in the case of S strain and R strain is virulent. What is mean by virulent is that it is not going to cause the disease whereas in the case of S strain it is virulent. So, it is actually going to cause the disease and what disease it is going to cause? It is going to cause the pneumonia and ultimately it is actually going to cause the death of mice. So, this is the experiment what he has done.

So, he has done the experiment using the four different types of conditions. So, case 1, case 2, case 3, case 4. So, in the case 1 what he has done is he has taken the S type of bacteria and that he has injected into a healthy mice and after some time the healthy mice is gone because he developed the disease pneumonia he has developed the pneumonia and that is how he actually died. And the case 2 he has injected the R type of bacteria.

So, R type of bacteria is virulent bacteria.

So, it is actually not going to cause any kind of disease and that is how this particular mice is actually going to grow and that is how the mice is actually going to live. Then the third experiment what he has done is he has actually denatured the S type of bacteria. So, he has heat killed the S type of bacteria and then he injected that into a healthy mice and that also has not caused any disease and that is how the mice is survived. Then the third is he has taken a heat killed S type of bacteria. So, a virulent bacteria he has heat killed.

So, this is not live bacteria and then he added the R type bacteria which is live. So, this is live this is dead bacteria and then he mixed them together and in the case 4 and then he injected them into a healthy mice and when he done that he actually found that the mice have developed the disease and he they also died. So, this was very interesting that something which was not infective because in the case 2 you see that the R stain is not killing, but in the case of in the case 4 the R stain is killing the mice if it is being mixed with the heat killed S2 or S3 actually. So, these are the 4 different types of conditions and then what he has done is he has isolated the blood. He isolated the blood from this case number 4 and what he found that there was no R bacteria present it was all the S type of bacteria what is been isolated.

So, all the R type of bacteria is actually being converted into the S type bacteria in the case 4 and that is how he said that the material or the material which actually carries the information from this bacteria gone into this material gone from the heat killed bacteria into R type bacteria and that is how the R type bacteria got converted into S type bacteria. So, there is a and then that is that is how he actually pointed the term that this phenomena is going to be called as transformations. So, when the Griffith has injected a mixture of heat killed and live bacteria the mice died and living living S bacteria recovered from the dead mice. So, remember that he actually he has killed the he has actually injected a heat killed S type bacteria that he got the live S type bacteria because the genetic material what are present in the S type bacteria heat killed bacteria were gone into the R type bacteria and that has convert or that has actually started expressing its own genome and that is how it is actually going to cause the generation of S type bacteria. So, Griffith concluded that the R type bacteria had somehow transformed by the heat killed strain bacteria some transforming principle transferred from the heat killed S strain had enabled the R strain to synthesize a smooth polysaccharide code and become the S strain this must be due to the transfer of the genetic material from the S type bacteria to the R type bacteria.

However, the biochemical nature of genetic material was not defined from this

particular type of experiment. So, we do not know still that what is actually being transferred whether the DNA or protein or RNA. So, then we did they did the further more experiments. So, in a meantime, the Oswald and McCarty and McElroy actually did the more specific experiment to ask whether the factors which are actually converted the R type bacteria into S type bacteria whether it is the DNA protein or RNA. So, it was believed that the genetic material was made up of the protein they work on the transforming principle in Griffith experiment to investigated its biochemical makeup.

So, to determine the biochemicals from the heat killed S type bacteria which could convert the live R type bacteria into S cell they isolated the biochemical like they isolated the proteins DNA and RNA from the S cell bacteria and finally, found that the DNA from the S bacteria was needed to convert the R type bacteria. So, what they have done and what they have used for this particular type of experiment they have used a three different types of enzymes they have used a DNases. So, the function of DNA is that it is actually going to digest the DNA. So, it is actually going to destroy DNA. So, if you take the DNase and if you take any reactions, it is actually going to destroy the DNA.

So, it is actually going to remove the DNA part. Similarly, they have the RNase. So, RNase the function is that it is actually going to digest the DNA which means it is actually going to destroy RNA part or I will say it is actually going to remove the RNA from the reactions. Similarly, you have the another enzyme which is called as proteases and the target substrate for the proteases is that it is actually going to digest the proteins and it is actually going to destroy the proteins because it is actually going to convert the protein into the amino acids. This means it is actually going to remove the protein components means if you have a reaction for example, if you have a reaction and if you treat this with the protease for example, so then this protein this reaction is going to be a reaction minus protein because you have treated with the protease and that is all you have removed the protein. So these are the some handy enzymes what these people have used to answer the questions which one which biomolecule is actually the genetic material.

So what they have done they have repeated the same experiment what the Griffith has done that where you have the heat killed S type bacteria and then you have the live R2 bacteria and then when they were taking the heat killed bacteria they have treated that heat killed bacteria with protease. So they have either treated it with the protease or treated it with the RNAs or treated with Ross okay, so, in these cases what you have done is you have removed the protein in this reaction you have removed the protein in this reaction you have removed the RNA and in this reaction you remove the DNA. This means in this reaction you have D protein and RNA present in this reaction, you have the

DNA and protein present. So what are the things present you have the DNA and RNA present, right? You have because the protein is already been removed because of you have treated with protease. Then in this one since you are removing the RNA, you are going to have the DNA and protein and in this one you are going to have the RNA and the protein because you have removed the DNA, right? And then he tested which in which condition the mouse is actually going to die.

So in this case when you have no protein you have no protein mice is dying, right? This means the protein is not responsible for converting the R bacteria into a rulant S bacteria, right? Similarly, when you have no RNA, right? Then also the bacteria of mice is dying because it is still the R type bacteria can be getting transformed into S type bacteria. And the third is when you would not have DNA, right? So when you remove the DNA, you are actually going to see that you are having a healthy mice, right? This means in this one you have the RNA and protein but still mouse is not getting the disease which means R strain is not getting converted into S strain and that is how the mouse is actually healthy. And what you see it even at the cultural level also when they culture the bacteria what they found that they could be able to recover the live S bacteria in the absence of protein and RNA but they could not be able to isolate the live S bacteria instead they are actually getting the R type bacteria R2 in the case of the third condition when the DNA is not present. So this actually confirmatory we proved that DNA is the genetic material.

Now let us talk about the conclusions. So they found that both RNA and protein digesting enzyme had no effect on to the transformation proving that the molecule undergoing transformation were either protein or RNA, okay? Or neither a protein or RNA transformation was prevented by DNA digestion indicating that the DNA was the transformation is in. So they came to the conclusion that the DNA is the genetic material but not all biologists agreed with that, okay? And then the further experiment were done by the Alfred Hershey and Cshershy and Chase experiment. So in the Hershey Chase experiment the researchers on the virus that infect the E. coli provided additional evidence for the genetic importance of the DNA. The DNA core of the T2 bacteriophage is in placed in a protein code.

Alfred Hershey's and M.R.C. Chase put this theory to the test in the following manner. The radioactivity labeled T2 phase in either the protein which means the 35S or they have labeled the DNA with the help of 32P component before injecting them into the bacteria which means they are going to have the bacteriophage where they have labeled the protein or where they have labeled the DNA by the radioactivity. Similarly they infected the non radioactive E. coli with the radio labeled T2 bacteriophage. So in a T2 bacteriophage you can have the protein which is 35S which means sulphur labeled or you can have the DNA which is actually going to be the 32P right or the phosphate

labeled, okay? And that is how they have asked where or which molecule is going from one generation to another So they injected the bacteriophage bacteria with the T2 phase that has been radio labeled either in the DNA component or in a protein component.

The infected bacteria were agitated in a blender and the two fraction were separated by the centrifugations. One fraction contains the empty fast code that was released from the surface of the bacteria that consists of the protein and therefore carries the 35S radio label. The other fraction consists of the infected bacteria itself. Most of the 32P label was found into the infected bacteria which means when they were doing the agitation what they found is that the protein which is a part of the code was always been extracellular and it was not being carried within the bacteria right? It was not getting into the bacteria and so when the virus is infecting the it is code is remain outside and that is how it is actually going to be present outside. This means the radio labeled protein remain outside whereas radio labeled DNA remain inside.

So that is why there was no sulfur what has been associated with the bacterial cell whereas in the case of DNA all the DNA was present inside and that is why there after centrifugations the radio activity was associated with the bacteria. So by doing this experiment where they were doing the infection followed by blending and followed by centrifugations what they found is that the protein is extracellular whereas the DNA was associated with the bacterial system and by doing this experiment the Hershey and Chase concluded that the DNA is the genetic material. So by the observation they found that the most radioactive protein was released into the supernatant whereas 32P DNA remained within the bacteria. Since genetic material was injected and T2 progeny was produced DNA must have been carrying the genetic information for the T2. This means not only that when the bacteria got lysed it has reduced in homogeneous amount of viruses.

This means the genetic information was there inside the bacteria to produce the viruses and that was nothing but the DNA. As host or the code of the bacteriophages were not labeled with the 32P and only with 32S the result of the experiment clearly indicate that the only DNA and not the proteins entered the bacterial cell. Protein code is left outside all of the genetic data necessary for the creation of a new fast particle is carried by the DNA that entered the host cell. This undoubtedly demonstrated that the DNA not the protein serves as the bacteriophages genetic material. And that is how they have concluded that the genetic material is DNA not the protein.

Then there are several indirect evidences for DNA to be as genetic material. First evidence is that the DNA regularly present in the nuclei of all cell types. It is equal the amount of DNA present in all cells of an organisms and amount of DNA is proportional

to the ploidy of the cell. The ploid cells have the half amount of DNA than the diploid cells. Nuclear division occurs only after the DNA duplication during the S phase of the interface is anyway we have discussed when you are discussing about the cell division.

Then the different species have different amount of diploidy DNA. Out of all macromolecule DNA is metabolically more stable or the most stable molecule and that is a first criteria that the genetic material should be very stable. Indefinite number of combinations are possible with the four sub bases like ATGC. DNA has some same physical and chemical property in all organisms yet allowed to produce great diversity of the organisms. So this is was clear that the DNA is the genetic material. But this has been challenged when people have discovered that there is a phenomena which is called as the reverse transcriptase or reverse transcription.

When the people have discovered a phenomena which is called as reverse transcription. So what is mean by the reverse transcription is that the RNA is actually going to give rise to the DNA. So this reverse transcription was against the central dogma of molecular biology. It says that RNA can be able to produce the DNA and by doing so there was people who said that RNA is also made up of is very stable right RNA is also stable RNA is also can provide the diversity and since RNA can be converted into DNA is there is a possibility that the RNA would also be behaved like a genetic material right and then the same in under the exceptional cases or some other kinds of cases right. So then again the same debate started whether RNA could be a genetic material or not.

So to prove that the people have started doing this experiment okay. So RNA as a genetic material okay. So according to the RNA world hypothesis the RNA was the first genetic material that stored all genetic information and it is believed that the first life arose from it. RNA is thought to catalyze a number of chemical reaction in the primitive cell. The presence of the two hydroxyl group in ribose group increase their reactivity but this reactivity makes them unstable which makes the RNA unfavorable as a genetic material. So as a genetic material it should be well stable chemically and structurally.

So that was one of the drawback of RNA as a genetic material that it actually contains the two prime hydroxyl group and because of that it is actually having the more reactivity compared to DNA. So ultimately these unstable molecules are replaced by the more stable genetic molecules. During this stage of evolution the DNA molecule emerged they have replaced RNA's role in the cell as both genetic material and structural component. The unstable and degraded nature of RNA has led to the development of double standard DNA genetic material that is both chemically and structurally more stable. So according to the hypothesis it is found that the RNA is actually being the preferred material for genetic material in the primitive cell okay.

So during the evolution what the cell has found that RNA is good in terms of genetic material because it reduces the steps right. You do not have to go for the transcription you can directly use that particular information to produce the protein. But on the other hand it is unstable so what they have done is they have converted the RNA into a double standard DNA and that is how you are actually bringing the more stability and chemically as well as structural stability into the structure. However RNA is not being completely eliminated they still serve as a genetic material in some systems like the viruses and they catalyze few essential biochemical reaction into the cell. Also the complex machinery of protein synthesis from DNA is still profiting through RNA okay.

So RNA is very important in terms of relaying the information from the DNA and that is why it is not been excluded from the complete picture of the protein synthesis. Still the protein is been synthesized from the RNA cell. So even if you see that the DNA actually stored the information it is RNA who actually you know dictate the production of the protein and that is how RNA is actually responsible for the particular type of phenotype. So RNA is present as a genetic material in some of the viruses right. So virus consist of two parts nucleic acid and the protein coat sometimes with additional envelopes.

So virus contains only one type of nucleic acid either the DNA or the RNA. These have RNA are called riboviruses they vary in the structure of their nucleic acid also the plant viruses are RNA viruses either single standard or the double standard. So you can have the animal viruses you can have plant viruses you can have single standard RNA viruses you can have double standard RNA viruses and in even in the plant viruses also you can have the single standard viruses like the TMB or you can have double standard viruses like the orayas viruses. Then we have the evidence in favor of RNA as a genetic material. So first evidence that RNA is also has a capacity to carry the genetic information came from the experiment which is conducted on the tobacco mosaic or TMB viruses. So Cricker and Schumann demonstrated in 1956 that the tobacco plant can contract mosaic disease when exposed directly to pure RNA from TMB.

RNA treatment render the pure RNA incapable of inducing the TMB lesions. Then Franklin, Pernhart and Singer demonstrated in 1957 that the progeny viruses from the TMB infection with viruses have RNA from one strain and protein from another strain were invariably of kind determined by the RNA not the protein. So this is what exactly they have done they have taken a TMB virus and what they have done is they have removed the capsid protein so they have removed fragmented that into a protein and RNA and then they degraded the RNA with the help of the RNA. So these are the four components so when they take the TMB virus and if they infect the new cells new leaf they could found that the infection is happening. When they remove the capsid okay

when they remove the cap a capsid part right so still it is actually having the RNA right and if you remove the cap capsid part or it is so if they remove the fragmented fractionated that into a capsid and the pro and RNA when they taken the capsid protein and infected that to the protein there is no infection so there is no infection right no infection. But they have take the RNA and if that infected into the protein and into the leaf what they found is there is a infection.

So basically the genetic information what was present into the RNA is good enough to produce the virus and that is how it actually causes the disease. Then they degraded the RNA with the help of the enzyme RNAs so there is no RNA right. They could found that there is no infection so there is no infection in this case also. So Grigor and Schumann correctly concluded that the viral genome of TMB is composed of RNA so if there is no RNA then there will be no infection. And then Franklin, Godhart and Singel uses the type A and type B TMB virus in this investigation the RNAs were then isolated RNA for the protein porting and then in order to create hybrid viruses Singel concluded the RNA of the strain with the protein of the another.

So the phenotypically and genotypically identical progeny virus was similar to the parental type from which the RNA had been recovered after rubbing the hybridization or reconstituted viruses on to the living one. So these people what they have done is they have excluded the RNA of they have isolated the RNA of the from the two different types of TMB viruses like TMB A and TMB B and when they mix them together what they could found that they are actually having the you know the hybrid viruses what is being produced. So in conclusion so Franklin, Godhart therefore come to the conclusion that both the DNA and RNA can carry the kinetic information as a result of all these investigations. His research established that the genetic material for TMB is stored in RNA rather than protein and however the DNA may always serve as a genetic material but DNA RNA is typically non-genetic RNA only serve as a genetic material in few instances when DNA is not present. So in summary what we have discussed we had discussed that the gene carries the data for the phenotypic expression and these genes are being called as a factor in the case of Mendel and the chromosome have the gene on them they are made up of the 60% protein and 40% DNA and stable genetic material that can replicate store information for expression and undergo the mutation is required and experiments by the Griffith every Hershey and Chase have reduced results that directly supports the DNA as a genetic material there are also some circumstantial arguments in favor of the DNA as a genetic material.

So experiment on TMB by the Grigor and Schumann demonstrated in 1956 that the it is the RNA which is actually carrying the information from one generation to another generation and that's how wherever you have the RNA it is actually going to be

responsible for generation of or causing the disease. Even when they have mixed the two different types of viruses what they found is that they are actually generating the hybrid viruses and the protein has no role in carrying the information from one generation to another generation. So by doing all these experiments it is concluded that the RNA is not being preferred as the genetic material until the DNA is present but if the DNA is absent the RNA is also being taken up as a substitute for carrying the genetic information from one generation to another generations or I will say that RNA is being preferred by the primitive organisms or such as viruses and whereas DNA is because DNA is more evolved so DNA is then taken up by the higher organisms and that's why very classific very categorically you can see that RNA is one of most referred genome in the case of the viruses whereas DNA is more preferred in the case of higher animals. So with this brief discussion about the genetic material we would like to conclude our lecture here in our subsequent lecture we are going to discuss some more properties of the genetic material and how it actually has a role in synthesizing the protein and we also going to discuss about the central dogma of molecular biology. So with this I would like to conclude my lecture. Thank you. .

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 04
Central Dogma of Molecular Biology
Lecture-18 Genetic Material (Part 2)

Hello, everyone. This is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering, IIT Guwahati and what we were discussing, we were discussing about the different properties in the course molecular biology. So far what we have discussed, we have discussed about the basic properties of the cells and then we have in the previous module we were discussing about the different types of biomolecules and we have discussed about DNA, we have discussed about the proteins, we have discussed about the enzymes and in this content in the current module we were discussing about the genomic DNA or the genomic genetic material actually. So, if you recall in the previous lecture we have discussed about the different types of experiments, how the people have figured out that which biomolecule has the potential to carry the information from the one generation to the next generation and continuing this discussion we are now going to discuss more about the genetic material and what is the makeup of genetic material and how the genetic material is actually going to be packed within the cell. So, in today's lecture we are going to discuss about the genetic material and how it is actually going to be packed into the tiny structure which is called as nucleus in the case of eukaryotic cell and in the case of prokaryotic cell it is going to be packed into a non-nucleus structure also.

So, when we talk about the genome or when we talk about the genetic material. So, the first question comes what is the genetic material? So, genetic material is a complete set of DNA comprising of nuclear and mitochondrial DNA in an organism and that is mean collectively mean called as the genetic material. This is definitely not the acceptable definitions as far as the prokaryotic is concerned because the prokaryotic cells does not contain the mitochondria. So, we are actually going to discuss about the prokaryotic structures, but in general the most acceptable definition about the genome is that it is a complete set of DNA comprising of the nuclear and as well as the mitochondrial DNA which means it is actually the DNA or it is actually the complete set of DNA which is going to be present inside the particular type of cell irrespective of whether it is a prokaryotic cell or a eukaryotic cell.

It is hereditary material which is present in an organism. So, the main purpose of the genome is that it is actually going to carry the information from the one generation to the next generation. If you recall in our previous lecture we discussed about the different

types of traits and so on although this particular course is not allowing us to discuss about the Mendel's experiments and we are not discussing about the genetic information how it is flowing from one generation to the next generation what are the different laws which are governing those kind of you know the movement of the genetic material from one generation to the next generation and so on. But it is actually the hereditary material what is present inside an organism irrespective of whether it is a prokaryotic cell or the eukaryotic cell. In the previous lecture if you recall we said that right it could be a DNA or it could be RNA.

So, because in the case of so many organisms it is could be a DNA or it could be RNA. So, the genome is totality of chromosome unique to a particular organism or any cell within the organisms. Each genome contains all of the information needed to build and maintain that particular organism. So, this is also very very important points actually there are two important point here one is that it is actually hereditary material what is present inside an organism and the second is that it is actually going to contain all the information needed for an organism to build and maintain that organism which means it is actually going to have all the informations of even about the developmental stages also how the person is how the organism will go through with the different developmental stages it is actually going to have those kind of information also. So that the organism is actually going to have the required changes in the body and so on.

For example, in the humans right in humans you are actually going to born as a baby and then you are going to have the different stages of the developmental stages and then you will reach to the you know reach to the puberty and then post puberty you are going to be adult. So, and even before birth also there are so many developmental stages you are actually going to go through and all these developmental stages are completely been governed by the genome what is present inside the organisms. Now the first question comes is how the genome and the genotype is differs right. So, there are many times the student get confused whether what is the genome and what is genotype right. So, genome is actually the hereditary material or the total hereditary material what is present in an organism is called as genome whereas, a part of the genome is actually been called as genotype.

The information contained within the chromosome or I will say a part of the chromosome this actually been called as the genotype. For example, you can have the genotype for tallness you can have the genotype for dark skins you can have the genotype for other brown eyes you can have the genotype for grey hairs and so on. So, these are the some of the properties which are actually going to be localized within a small portion of the genome they are not going to be completely there will be a genome for that particular thing right. So, genome is a collection of genotypes and the genotype

is a subset of that particular genome. Now, the question comes what are the different types of genome what are present in the different organisms.

So, we have the four categories of the genome which is according to the organisms you can have the prokaryotic organisms such as the bacteria you can have the eukaryotic organisms such as the animal and this is plant and then you can also have the specialized type of organisms such as the virus ok. So, in the case of bacteria or in general though prokaryotic organisms you can have the prokaryotic genome which is actually going to be the double standard DNA circular chromosome and then it also going to have the nucleoid. All these we are going to discuss in detail whereas, in the case of eukaryotic genome which is present in the animal you are going to have the double standard DNA you are going to have the linear DNA and it actually going to be present in the form of many chromosomes and all these are actually going to be present inside a confined structure which is called as the nucleus. So, this is the nucleus what is present right and then in the case of plants you are going to have the main genome what is present inside the nucleus just like as the animal cell and then you are also going to have the organellar genome which is going to be present in the mitochondria and the chloroplast. And then we have the viruses so, we can have the viral genome.

So, viral genome could be single standard DNA double standard DNA or it could be RNA it could be circular or it could be linear then it is segmented or non-segmented and then it is actually going to be monopartite or the multipartite. So, all these are the summary of the properties of these genome and the genome are actually been organized inside a particular organism right. For example, in the prokaryotic cell the genome is going to be distributed or going to be present within the cytosol whereas, in the case of eukaryotic cell the genome is either going to be present inside the organelles or it is actually going to be present inside a well defined structure which is called as nucleus. So, how the genome is organized? So, genome genetic organization is like that. So, in the cell the each DNA molecule associated with the protein molecules and the each DNA molecule and its associated protein is called as the chromosome.

So, in any organisms the DNA is actually going to be get associated with the protein molecules and this particular structure is going to be called as chromosomes. This organization is actually going to be hold true for prokaryotic cell or the eukaryotic cell. In the eukaryotic cell you are going to have the many types of chromosomes and the prokaryotic cell the genome is very small. So, it is actually going to have the single chromosome. So, this is just a classical example right you are going to have the eukaryotic cell you are going to have the nucleus within the nucleus you are going to have the DNA right.

So, this is the DNA right and then this DNA is actually going to be associated with the different types of proteins and that is how the DNA is actually going to be condensed in the form of different condensation organizations levels and then ultimately it is actually going to form the chromosomes. This is the similar kind of organization even in the prokaryotes the only difference is that the prokaryotes will have only one chromosome which is a circular chromosome. And the packaging of DNA into the chromosomes. So, the DNA is actually going to be packed into a dense material and that is going to be called as chromosome. The chromosome is actually going to have the DNA and it also going to have the protein and protein is actually playing a very crucial role in packaging the DNA into the form of chromosomes.

So, the packaging of DNA into the chromosomes serves several important features. Chromosome is compact form of DNA that readily fits into the cell right this is anyway we are going to discuss in detail. Then it protects the DNA from the damages and it only packed DNA can be transmitted efficiently to go to the daughter cell when a cell divides. So, since it is a packet you can actually be able to share these packets between the daughter cells very precisely. If it is a loose DNA then it could be possibility there is a possibility that you may actually share the 50 percent, 75 percent, 80 percent like that.

But if it is a packet you will either share the complete packet or you will not share the packet. So, that actually gives a regular flexibility as well as in the regulation that the DNA is actually going to pack into the form of a chromosome. Now the question comes why there is a need to pack the DNA. So, why the packaging of DNA is required? So, DNA is packaged into a form of chromosome and then these packaged DNA is actually going to be required for many reasons. Number one it is actually going to be required for DNA competition, it is actually been required for the DNA protections, it is actually been required for regulation of the gene expression, it is also required for facilitating the DNA replication and repair.

Then also ensure the accurate chromosomal segregations and then lastly it is also required for enabling the regulatory interactions. So, what is mean by the DNA compactation? So, DNA competition by packaging the DNA into the compact structures such as the nucleosome and the higher order chromatin fibers, the physical size of the DNA molecule is reduced significantly. You know that for example, the human genome right, human genome is approximately been of a size of 1 meter fiber right. So, if you have a 1 meter fiber and you know the size of the cell right, cell is approximately 30 micrometer right. So, if you have a cell of 30 micrometer and if you have a genome of 1 meter fiber, it cannot fit into this right, it cannot be fit into this.

So, to fit this you are actually required to compact this to pack it such a dense material

that it should actually fit into this particular site. So, that is the purpose of packaging the DNA into the chromosomes. Then the second is that it is actually going to provide the protection to the DNA. So, the number 2 is it is actually going to provide the protection into the DNA. So, the densely packed chromatin structure shielded the DNA from the exposure to the potentially harmful agents such as chemicals, radiations and enzyme.

It also help to prevent the DNA from becoming tangled or breaking during the cellular processes. So, you can imagine that if I have a DNA which is loose DNA right, if it is a DNA which is loose it is actually been accessible for all sort of damaging material. For example, if you are taking a you know if you are getting exposed to the free radicals or if you are getting exposed to the hydrogen peroxide, it is actually going to have the direct access to the DNA. And this anyway we are going to discuss in our subsequent module when we are going to discuss about the DNA damage and repair. But so, there are several type of DNA damaging agents right, one is the free radicals, the other could be alkylating agents, the drugs what you are taking and all that.

So, if you are taking a drug and if the DNA is not properly packed, it will getting exposed and that is how it is actually going to be get damaged right. It is going to be damaged because there are you know modifications in the nucleotide and so on. This is completely been protected by if you have the DNA and if this DNA is been surrounded by the protein molecules. So, now what you have is if you have even these molecules, they will actually going to interact with the protein rather than the DNA. So, the drug will go and interact with the DNA because if the DNA is not going to be accessible because of it is surrounded by the different types of proteins and that is how the DNA is getting protected.

Now, the third point is the regulation of the gene expressions right. You know that the gene expression is a very tightly regulated process right. So, if the gene regulation is not been done, then it is actually going to have a very very significant negative effects on the health of that particular cell. For example, you have taken the food right and you have taken the meal right, you have taken the food, it has produced the glucose right or and the glucose the blood it has increased the blood glucose level. Now, if I have to tackle this problem, what I have to do is I have to secrete and I have to synthesize a large quantity of insulin right.

That means, as soon as this occurs, I have to do a gene expression profiling, I have to change the gene expression profiling within the pancreas and within the pancreas and as a result what will happen is the pancreatic beta cell are actually going to start secreting the insulin. That is actually going to affect on to the some of the effector organs like such as liver and muscles and that is how they will actually going to convert the glucose into

the glycogen right and that is how they are actually going to protect the body from the harmful effects of having the very high level of blood glucose. Now, this is temporary right, this effect is temporary because after some time the blood glucose level will reach to a normal level and then if this process will continue, then it will actually going to go down to the liver right. For example, if the blood glucose level is 80 milligrams per deciliter right, which means 80 milligrams per 100 ml, then it is the normal level right. But when as soon as you have taken the food, the level will go down go up right, level will go up to like for example 200 right and from 200 it will return back to 80.

But if this process will continue, it will further come down right, it will come down to 50 right and it will come down to 0 if this will continue because the insulin does not know that there is a glucose, there is enough amount of glucose right. So then so there is a regulation of gene expression required right, as soon as this reaches to 80, then the blood glucose level or it is actually going to give the indication to the pancreatic cell that ok, there is no more insulin required and that is how it is actually going to change again, it is actually going to change the gene expression profiling and that is how there will be no secretion of insulin. And that is why it is very important for maintaining the normal physiology of an organisms. So regulation of gene expression, the DNA packaging in can influence the accessibility of the gene to the cellular machinery involved in the gene expression, such as the transcription factor and RNA polymerase. By compacting or loosening the chromatin structure, cells can control which genes are accessible for transcription and thus regulate the gene expression pattern.

So we have discussed how the gene expression profiling is going to have the significant effect on to the overall physiology of that particular organism and it is been influenced completely by whether the DNA is present in a compact structure or not. Because if it is if the gene is present in a compact structure and it is not accessible for the cellular machinery to perform the transcription and translation, then that will not going to be transcribed. On the other hand, as soon as you would like to have the down regulation of a particular gene expression, you just put that particular gene into a tight compact structure and that is how it is actually going to control the overall gene expression of that particular protein. Now, the fourth point is it is actually going to facilitate the DNA replication and repair. This anyway we are going to discuss in detail when we are going to talk about the DNA damage and repair and all that and replication also.

So during the DNA replication and repair process, the packaging of DNA into the nucleosome must be temporarily loosened to allow the necessary protein and enzyme to access the DNA strands. After replication or repair, DNA is repacked into the nucleosome and hired or chromatin structures. Hence, the proper genome packaging ensure accurate replication and repair of DNA. So this is very very important that we

should have the packaging and the unpackaging of the chromosomes so that some amount of DNA is going to be open and then that DNA is going to be replicated.

Now the number of 0.5 is it is actually going to ensure the accurate chromatin segregations. So during the cell division, the genome must be accurately divided between the daughter cells, the compact organization of the DNA into the chromosome facilitate this process. This anyway we have discussed in detail that if you have a single chromosome, you will divide this chromosome and you will make two chromosomes. And then you will actually going to divide this chromosomes equally. You are going to take give one to the sister or the daughter cell and one you are going to keep it along with the parents.

If it is not compact, if it is a DNA, if it is a DNA, if it is not compact, if it is a DNA, then the division could be 70, 80 or percent. It could vary because some amount of DNA you will put it into a sister or daughter cell, some amount of DNA you will put it into parent cell and so on. So there are possibility that the daughter cells will have 1.25 copies of the genome and the parents will only left with the 75 percent. So this kind of possibility should not exist that is why the DNA is actually going to be packed into a packets.

So you will just take one packet with you among yourself and then you are actually going to give one packet to the daughter cell and that is how there will be an equal division of the genetic material between the two daughter cells. Then we have the enables the regulatory interactions which is the point next point. The 3D organization of the genome within the nucleus allow for the specific regulatory interaction between the different region of DNA. This is special organization facilitate the long range interaction between the enhancer region of DNA. So this is all about why the DNA is required to be packed into a compact structure to form the chromosomes.

Now let us see how you can be having this kind of organized this kind of packaging into the different types of organisms. So we will start with the prokaryotic organisms. So first we are going to discuss how the genome is going to be organized into the prokaryotic structures. So the genome organization in a prokaryotic cell. So they these are this is the typical prokaryotic cell where you have the cell wall you are going to have the capsule you are going to have the plasma membrane and inside the plasma membrane you are going to have the cytosol and then you are going to have the nucleoid.

So nucleoid is the region where you the all the genomic content is going to be present and then you also have the ribosomes and the plasmids and you are going to have the all this kind of phagella and all that. So they have these they have these small bodies and a

small genome. You know that the genome bacterial is bacteria cells are very small. So they are going to have the smaller genomes. They do not contain a nucleus or any other membrane bound organ.

This all we have discussed when we were talking about the cellular structures. Then they have a small circular DNA which is present inside the nucleoid region. So this is the nucleoid region where the genomic content of the prokaryotic cell is going to be present. They have a single chromosome that floats within the cytoplasm. The genome size ranges between $10 \text{ to power } 4$ to $10 \text{ to power } 7$ base pair with a high gene density.

Apart from this single chromosome some bacteria have the extra chromosomal DNA which is called as the plasmids. We have also discussed about the plasmids when we were talking about the bacterial cell and we have also discussed how you can be able to isolate the plasmids from the bacterial cell and we have shown you a demo video also how to do that. So in a prokaryotic cell you going to have the chromosomal DNA as the genomic content. The second is you are going to have the extra chromosomal DNA which is also been called as plasmids.

So let us discuss about the plasmids. So prokaryotic frequently carry one or more smaller independent extra chromosomal DNA which is called as plasmids. Plasmids are not genomic DNA they are accessory DNA molecules. A smaller circular DNA molecule that have the ability to self replicate. Unlike the larger chromosomal DNA plasmids typically are not essential for the bacterial growth. So plasmids are actually been required for providing the specific properties and by the plasmids they can be able to exchange that property between the particular bacterial colony.

For example there could be a property like resistance against ampicillin. So if a bacterial colony has acquired a resistance against the ampicillin and that property it has captured in the form of a plasmid then it actually can share that plasmid within the colony and that is how they can be able to distribute that particular property within the organisms. What is the importance of the plasmids? So plasmids provide advantage to the bacteria such as antibiotic resistance, herbicide resistance etc. So all these properties are actually been due to the different types of antibiotic, different types of genes and all these genes are actually going to be cloned within the plasmids and that is how the plasmids are actually going to express the protein and provide the necessarily resistance mechanisms within the cell. And that is how the bacteria can actually provide that particular plasmid to the colony and that is how the colony is actually going to be resistance.

In addition unlike chromosomal DNA plasmids are often present in many complete

copies per cell. So unlike the chromosomal DNA the plasmid will not have the bacteria will not have the single colony single copy of the plasmid it could have the 200 copies it could have 500 copies and so on because the number of copies will decide which bacteria will have the higher resistance property. So if you have a very high number of plasmids you can be able to have the higher resistance for that particular antibiotic or that particular type of phenotype. Then we will talk about the bacterial genome.

So bacterial genome is very small. So bacterial chromosomal DNA is usually a circular molecule that has a few million nucleotides in length. For example in the case of E. coli you have the 4.6 million base pairs. Similarly you have the H influenza H influenza is going to have 1.

8 million copies. So it is actually a small genome what is present in the bacteria. And then a typical bacterial chromosome contain a few thousands different types of genes. So the unlike the eukaryotic organisms you are not going to have the useless genes you are not going to have the other kinds of non intron or non expression genes actually. So bacteria only contains that gene which are going to be expressed and which are actually going to have some meaningful effect or meaningful purpose inside the cell because you know that they have their size is very small so they do not want to cover or they do not want to keep the unwanted materials. Then structural gene sequence account for the majority of the bacterial DNA or the encoding proteins.

The untranscribed DNA between the adjacent genes are termed as intergenic regions and these process these regions are very very small or almost absent in the bacterial system. Then the since you have the DNA you have to pack this DNA into form of a chromosome so that you can be able to make the compactation and make the structure very compact. So packaging of DNA so prokaryotic cell usually have a smaller genome that need to pack their DNA is still substantial you know that we bacteria are few mu micron in size so they are DNA size is also relatively very big so it has to be compacted. E. coli must pack its 1 mm chromosome into a cell that is only 1 micrometer in length.

It is less clear how the prokaryotic DNA compacted but it is actually going to be packed into a small structure or within the cell. So the region what it is actually going to pack the DNA is called as nucleoid. So nucleoid is a primitive nucleus or I will say it is actually a primitive nucleus except that it is not going to have the membrane. So it is not a membrane it is a region in which the bacteria is actually going to have the chromosomal DNA. A prokaryotic chromosome is circular and it is reside in a cell region called as the nucleoid.

Only one complete copy of their chromosome that is packed into a nucleoid. 80% DNA by mass can be unfolded by agent that act on RNA or the protein. The proteins responsible for condensation and maintaining the super coiled structure of the DNA have not been identified. So that is still unknown right how the different types of proteins are involved and what are the different types of proteins are involved into the making the structure very compact. The type of protein found in prokaryotic chromosome known as the nucleoid associated protein which is responsible for compacting the DNA into a chromosomal structures.

DNA determine the protein what protein and enzyme an organism can synthesize and therefore what chemical reaction it can be able to carry out. So what is the function of the genome whether in any organism that it is actually going to determine the proteome of that particular organisms. It is actually going to decide what are different types of proteins and enzyme are going to be produced and that is how it is actually going to eventually we decide the metabolism and physiology of that particular organism. So this is the micrograph of a bacterial cell where the nucleoid is actually going to be shown. So this is actually the region where the nucleoid is present and within the nucleoid what you are going to have is the bacterial chromosome which is super coiled actually.

What are the key features of the nucleoid is that most but not all bacterial species contain the circular chromosomal DNA. A typical chromosome is a few million base pair in length. Most bacterial species contain single type of chromosome but it may be present in multiple copies. Several thousand different genes are inter dispersed throughout the chromosome and one origin of replication is required to initiate the DNA replications. So this anyway we are going to discuss when you are going to discuss about the origin of applications.

So origin of applications in the case of the bacterial chromosome is single. So it is actually going to start here and it is actually going to go through and then it will come to here right and you know that it is actually going to go in both the direction. So this is going to be leading strand and this is going to be lagging strand and that is how it is actually going to produce two copies of the genome after one cycle right. One will which will come from this side and the second which will come from this side.

So second will come like this right. So that is why it is actually going to have one original copy and the one is replicated copy. The short repetitive sequence may be inter dispersed throughout the chromosomes. The chromosome DNA must be compacted about a thousand fold remember that 1 mm will actually need to be compacted within the 1 micrometer diameter. So it is actually has to be compacted around thousand folds. The formation of the loop domains so and the number of loop varies according to the size of

the bacterial chromosome and the species.

For example the E. coli has 50 to 100 with 4000 to 40,000 to 80,000 base pair of DNA in each. So you are going to have a circular chromosome and then this circular chromosome is actually going to be looped and it is actually going to compacted by doing this and that is how it is actually going to form the very strong structures. Well let us take an example of one bacterial chromosome where one bacterial species how the chromosome is occurring.

So in the case of E. coli the E. coli chromosome is compact one fifth of a volume right. The determinants of the nucleoid folding so negative super coiling by the topoisomerizes and the condensation by the attachment of the nucleotide structure proteins. The nucleoid is highly condensed during the rapid growth and RNA polymerase concentrate in the transcriptional loci and the RNA polymerase is distributed throughout the chromosome. So this is the bacterial cell where we have already discussed about the different types of parameters. So this is all about the genomic material and what we have discussed we have discussed about the importance of the genetic material and what are the different types of genetic material are present in the different types of organisms.

So we have taken an example of the perocratic structures then we have also discussed about the Eukaryotic structures and when we also discuss about the viruses. So in the case of prokaryotic you have the double standard or double standard DNA and it is going to be present in the cytosol in the form of circular chromosome and then apart from that is also going to have the extra chromosomal DNA in the form of plasmids. In the case of Eukaryotic cell you are going to have the either the genetic material present inside the nucleus or it is actually going to have the genetic material which is present inside the organelle such as the mitochondria or the chloroplast. In the case of viruses it can have the single standard DNA or the double standard DNA or the RNA and the viruses are unique organisms so they are actually going to have the different types of physiology and different types of manipulation of their genetic material. So and then lastly we have also discussed about the how the genome is organized within the prokaryotic structures and how the compactation is occurring inside the genomic inside the prokaryotic cell and so on.

So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss some more aspects related to the genomic material or the genomic content. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 04
Central Dogma of Molecular Biology
Lecture-19 Genetic Material (Part 3)

Hello everyone, this is Dr. Vishal Thevedi from department of biosciences and bioengineering IIT, Guwahati. And in the current module, we are discussing about the genetic material. What we have discussed so far, we have discussed about the discovery of genetic material by doing the different types of experiments. So, initially there were two prominent candidate for the genetic material, one is the nucleic acid, the second is the proteins. And there was long debate about which material is actually going to serve the purpose of having a potential to be genetic material.

We have discussed about the different types of experiments through which the scientists have, you know, proved that it is actually the genetic nucleic acid which is actually been the most acceptable material for genetic material. So within the nucleic acid, it could be DNA or the RNA. Mostly the cells are actually having the genetic material in the form of DNA, whereas in some of the organisms such as viruses, you also going to have the RNA as the genetic material. Then in the previous lecture, we have discussed about the genomic organizations and how the genetic material is actually going to be organized in the different types of organisms.

So, we have discussed about the organization of the genetic genome organization in the prokaryotic cell and we have discussed about how the different types of materials are required for genomic organization in the prokaryotes, what are the properties of the genome into the prokaryotic structures. We have taken an example of E. coli and how the compaction is happening inside the E. coli cell and so on. Now in today's lecture, we are going to discuss about the genomic organization into the eukaryotic cell and how the different types of chromosomes are being found.

So in the genomic organization, when we talk about the genomic organization in the eukaryotic cell, we are having the genomic genome which is present in two places in the eukaryotic cell. One is the nucleus and the second is the organelles. So if it is a plant cell, it is actually so within the organelle, you can have the two different types of organelle where the genomic content is going to be present. One is called as the mitochondria and the second is the chloroplast. So you can genomic genome is actually been organized within the eukaryotes in the nucleus.

So within the nucleus, you are going to have the different types of chromosomes and within the mitochondria also you are going to have the mitochondrial chromosome and the chloroplast chromosomes and summation of these DNA material is actually going to be called as the genome within the organisms. Now within the genome organization in the eukaryotes, so the genome eukaryotic genome is linear and confirmed with western crete double-clicked cell structure model. It is embedded in the nucleosome complex DNA and protein structure that packed together to form the chromosomes. Eukaryotic genome have unique feature of exon-intron organization of the protein coding genes representing the coding sequence and the intervening sequences that represent the functionality of the RNA inside the genome. One haploid cell consists of the 23 different chromosomes and one mitochondrial chromosomes contain more than 3.

2 billion DNA base pairs. Remember that in the case of the prokaryotic structure, we were talking about the number of chromosome in the case of millions, but here the number of chromosomes are going to be in millions that is why there is a more higher level of compaction required so that that particular DNA is going to be fit into the small tiny nucleus. Now as far as the chromosomes, so eukaryotic chromosomes are usually linear structure. So this is the eukaryotic one of the eukaryotic chromosomes and a typical chromosome is 10 of millions to hundreds of millions of bases where in length. Eukaryotic chromosomes occurs in set many species are deployed which means that the somatic cells contain two set chromosomes two sets of the chromosomes right.

So you are going to have the chromosomes which are present in the sets right and so double set is present inside the cell and during the duplications or during the division one set is actually going to be shared with the daughter cells. Each chromosome contains a centromere that forms the recognition site for the kinetochore complex. So this is the centromere right this is the centromere and this is the place which is actually going to be recognized by the kinetochore proteins. Then we have the telomeres which contain the specialized sequences located at the end of the linear chromosomes. So these are the telomeres which are going to be two telomeres.

So and this is the specialized sequences which are present on the tails or the corners of the chromosomes. Then we have the repetitive sequence which are found in near the centromeric regions. So these are the DNA or the genes how the genes are actually been organized onto the chromosomes. For example this is a classical example how the MSC proteins are or MSC genes are been organized onto the chromosomes within and you see that you have HLA typing and all that and this is a particular type of chromosomes. What is the chemical composition of the chromosome or the chromatin? So you are going to have the DNA you are going to have the RNA and then you also going to have the proteins.

So you are going to have the DNA which is 20 to 30 percent. So most important constituent of the chromatin then you are going to have the RNA which is going to be 5 to 10 percent. It is associated with the chromatin as the transfer RNA messenger RNA or ribosomal RNA. So it is actually be a part of the expression machinery and then you also going to have the proteins so that is going to be 50 to 60 percent. So mostly the chromatin is actually made up of the DNA plus protein and it is mostly the 40 percent is actually going to be DNA and 60 percent is actually going to be the protein.

Within the protein you are going to have the two different types of proteins you are going to have the histone proteins or the non histone proteins. So histone proteins very basic proteins constitute about 60 percent of the total protein almost it is present in the 1 is to 1 ratio with the DNA and you are going to have the five different types of histone proteins H1, H2A, H2B, H3 and H4. Similarly you are going to have the non histone proteins they are 20 percent of the total chromatin protein and they are required for the nucleosomal assembly proteins such as NAP and the other histone chromatin remodeling complexes and then you are going to have the structural proteins such as actin, tubulin and myosin and the contractile protein and all the enzymes. So the contribution of the non histone protein is very very little compared to the histone protein histone protein is required for packaging. So histone is a positively charged protein so these are the basic proteins which means they are actually going to be positively charged protein and that is why they are actually going to have the instant attraction for the negatively charged DNA.

So this is the DNA you know that the DNA is negatively charged because of the phosphate backbone. So the histones are found in all the eukaryotic cells they are commonly present in the histone commonly present histone in the eukaryotic cells are H1, H2A, H2B, H3 and H4. Then these all the five histones are being categorized into the two structural regions one is called as the core histone the other one is called as the linker histone. So the H2A, H2B, H3 and H4 are part of the core histone whereas the H1 is actually being called as the linker histone. So core histone the two copies of the core histone form the protein core around which the DNA is wrapped and within this you are going to have the H2A, H2B, H3 and H4 whereas the linker histone it is not the part of the core protein but it is associated with the linker DNA which links the two nucleoside.

So within the histone you are going to have the core histone and the linker histone. The core histone you are going to have the two copies of the H2A, H2B, H3 and H4 which are actually going to form the core of the nucleosome and then this core is actually going to wrap the DNA. So it is going to have the clear attraction because all these histones are going to be positively charged so their surface is actually going to be positively charged

and that is how they are actually going to have the very high affinity for the DNA. Similarly you are going to have the linker histone which is going to be H1. So how these histone proteins are or what are the different properties of the histone proteins.

So histones are closely associated with the negatively charged protein DNA. They have the high content of positively charged amino acids such as lysine and arginine. So you have the core histones and you have the linker histones. Core histone H2A you are going to have the molecular weight of 14 kilo Dalton and it is actually going to be a lysine rich. Similarly you can have the H2B is going to be approximately 14 kilo Dalton and it is the slightly lysine rich.

Then we have the H3, H3 is going to be 15 kilo Dalton and it is going to be lysine rich and you have the H4 which is going to be 11 KDA and it is going to be arginine rich. Then we have the linker histone which is H1 and it is going to be 20 KDA and it is also arginine rich. So H3 and H4 histones are first formed the heterodimer then come together to form a tetramer with the two other the molecule of H3 and H4. H2 and H2B form the heterodimers. So this is the sequence in which the histones are actually going to be organized with each other and that is how they are actually going to form the nucleosomal assembly.

So the assembly of nucleosome involved the ordered association of histone with the protein and DNA. So first the H3 the two molecule of H3 and the two molecule of H4 are actually going to come together and that is how they are actually going to form a tetramer which means you are going to have the like this. So you are going to have a tetramer which is going to be formed. Then this tetramer is actually going to because this tetramer is going to have the positive charge. So this is going to have a positive charge on the top right and then it is actually going to bind the DNA.

This means this ball is actually going to have the DNA. And then to after binding the DNA then the H2A and H2B is dimer which means is actually going to join the H3H4 DNA complex and that is how it is actually going to have the nucleosomal So this is actually one ball which is going to have the H3H4 DNA and on top of this you are going to have the H2A H2B and H2A H2B like that. So DNA is actually going to be present inside this particular core. So this one copy of nucleosome is actually going to communicate with another nucleosome right so this is going to have another copy and this is the region where the H1 is actually going to bind. So this is the lingered region which is actually going to bind.

The core histones have the amino acid extensions called as TALE because it lacks the defined structure and do not participate in association of DNA with the histone octamer.

The TALEs are the extensive site for a post translational modification including the methylation, acetylation and the phosphorylation. The assembly of the histone involves the ordered association of the histone with the DNA. This is anyway we have discussed right H3H4 followed by the binding of DNA followed by the binding of the H2A H2B and that is how the nucleosome is going to be formed. So the nucleosome is the starting building block for the higher order organizations.

So nucleosome is actually the building block of the chromosome right so this is the DNA right then it is actually going to form the nucleosome then it is actually going to organized into the chromosomes and then it is actually going to organize into the nucleus. So a human cell contains 3 into 10 to power 9 base pair per haploid set of the chromosome. The average thickness of each base pair is 3.4 angstrom this is the thickness of the DNA actually. So therefore if the DNA molecule in a haploid set of chromosome were lay out end to end the total length of the DNA molecule would be approximately

1

meter.

For a diploid set the length is actually the double which is the 2 meter as the diameter of a typical human nucleus is about 10 to 15 micrometer it is obvious that the DNA must be compacted by many order of magnitude to fit into such a small space. The compactation in human nucleus is done by the nucleosome formation by the association of the DNA with the histone. Nucleosomes are packed into successively high ordered structures. So nucleosome model is scientifically model which explains the organization of the DNA and associated protein into the chromosome. It also further explains the exact mechanism of the folding of the DNA into the nucleus.

This model was proposed by the Kornberg in 1974 and it is the most acceptable model of the chromatin organization. The model was further confirmed by the Pea audit in 1975. What are the features of the nucleosomal model? In eukaryotic DNA in eukaryotes the DNA is tightly bound to the histone protein which leads to the formation of DNA protein particles called nucleosome. Neustones play a very important role in packaging of such a long DNA molecule in the form of a nucleosome into the nucleus only a few micrometer in diameter. Therefore the nucleosomes are called fundamental packaging unit particle of chromatin and it gives a beads on to a string appearances which means if you see very clearly you will see that the DNA is being packed like this.

And this packaging is called as the beads on a string because this is the linker DNA and this is actually the core structure of the histone and that is how these are actually going to be you know fold on to each other. So it is actually going to be fold like this and that is how you are going to have a beads and then it is going to be fold like this and that is how it is actually going to be condensing and that is how you are going to have the higher

organizations of the packaging. Each nucleosome is a disk shaped particle with a diameter of 100 nanometer and 5.7 nanometer in height containing two copies of each four nucleosome histones such as H2A, H2B, H3 and H4. This histone octamer forms a protein core around which the double standard DNA is wrapped to 1.

6 times containing the 146 base pair. Each nucleosome bead is separated from the next by a linker DNA which is generally 54 base pair and contains the single H1 protein. This is what I was talking about right you have a DNA and then it is actually going to be arranged. On average the nucleosome repeat at an interval of 200 base pair. So folding of the DNA so once the nucleosome is formed there will be a folding of the DNA and that is how you are going to have the higher compaction of the DNA. The assembly of DNA begins with a newly produced tetramer which is H3H4 that is a particularly modified to form a sub nucleosomal particle the two Hb-HTM are then added.

Results in the formation of a nucleosomal core particle with a 146 pair DNA bound to the histone octamer the nucleosome is made up of this central component and the connecting DNA. In order to create the nucleo filament the nucleosome core must be spaced regularly which is accomplished during the maturation stage which requires the ATP. The newly integrated stones are deacetylated in this stage. Next the incorporation of linker stone is accomplished by folding of the nucleo filament into the 30 nanometer fiber the structure of which remain to be elucidated. Two principle model exist one is the solenoid model and the other is the zigzag model actually to explain how the DNA is actually going to be folded after forming the nucleosome to form the higher order organizations.

Last but not least further subsequent folding process results in a higher level of structure and distinct domain within the nucleus. So these are the some of the organizations so you are going to start with the DNA. DNA is actually going to be fold into the so if you start with the 2 nanometer then you are actually going to form the beads on a string form and that is how actually you are going to have the nucleosomes. Then these nucleosomes are actually going to be organized and refolded the DNA is actually going to be folded again and that is how you are going to have the 30 nanometer chromatin fibers. Then these 30 nanometer chromatin fibers are actually going to be further organized and they are actually going to form the 300 nanometer fibers and in which these solenoids are actually going to be again folded onto each other and that is how you are going to have this 300 nanometer fibers.

Then this 300 nanometer fiber is actually going to be condensed and that is how it is actually going to form the 1700 nanometer fiber and you are going to have the condensed region of the chromosome and from this you are going to have the further condensation

and as a result it is actually going to form the 1400 nanometer for chromosomes. So you are going to have the chromosomes so you started with DNA you ended with the chromosomes. So this is the first and this is the last binding style. Now in this is the packaging of the genome into the eukaryotic system and the genome has a very significant impact on the properties of that particular organisms and the bigger the genome you are actually going to have the more of the information what you are going to carry and that is how you are actually going to have the higher flexibility of modulating that information and that is how you are going to have the more properties to handle. So genome size is related to the complexity of that particular organisms because higher the information you have the higher you are actually going to manipulate that information and that is how you are going to have the complexity into the system.

This means you are going to you can actually have the freedom to modify the proteins you can actually be able to synthesize the proteins and so on. So if you see a very clearly what you are going to see is you are going to have the prokaryotic species and you are going to have eukaryotic species. Within the prokaryotic species you will see that the genome size is very small and that is why the number of genes or number of proteins are actually going to be very small. This means they are actually going to produce the less number of proteins and if you are producing the less number of proteins you are actually going to have the lower order freedoms to manipulate those proteins because you cannot have the multi step process you can have only the one or two step process because you if as many number of process you are going to produce you can actually should have the proteins to regulate these steps. So that is basically a drawback or I will say simplicity in the system right.

More and more you are actually going to have genome size for example in the case of yeast you are going to have the genome size which is 12 megabyte right and then you will see that the number of genes are going to be significantly very high compared to the mycoplasma and more and more actually you see that these are the plant species right. So they are very very high and the number of genes what you see here is very high this means they are actually being able to have the potential of producing the large number of proteins and that is how these large number of proteins can be utilized in such a way that you are actually going to have an event which is tightly regulated at each step and that is how you are actually going to have the you can actually be able to have the more control over the process. So the C value or the cot values or the quantity of DNA per haploid genome such as that seen in the nucleus of *S. permatozone* is used to describe the genome size in the eukaryotes because the size is essentially consistent within the species it is known as the C value or the characteristics. The mismatch between the C value and the presume amount of genetic information contained within the genome was called C value paradox.

Since we cannot assume that a species processed less the DNA than the quantity required to specify its vital function we have to explain why many species contain this amount of excess DNA. This is very simple actually if you have excess amount of DNA you can have the flexibility of producing more number of proteins and that is how you can have the instead of having the 3 step process you can have the 20 step process because if you increase the number of steps you are actually going to have the flexibility. You have seen that the glycolysis is a 10 step process. Glyceraldehyde cycle also has a multiple steps and because they have the multiple steps you can actually have the entry and exit of the metabolites at every stage and that is how you are actually going to have the very very complex biochemical reactions. Now the first question comes if there is a requirement of the protein production for example if there is a requirement of protein production that is required that the DNA should be free for doing the transcription and translation.

So now question comes how you can be able to unpack the DNA and how you can be able to have the DNA which is available for doing the other kinds of molecular biology activities such as replication, transcription and translations. So unpackaging of DNA so the way we have discussed about the packaging of the DNA the same way the histone has the crucial protein which are actually going to participate into the unpackaging of the DNA or the unwinding of the DNA. So histone are actually having the tail region right remember that tail region is important for histone to be assembled with each other and that is how they are actually going to bind the DNA. This tail region has the modification site for acetylations, for phosphorylations and the methylations. Now when you have the acetylation you are actually going to have you are going to produce the negative charge.

When you have the phosphorylation you are going to induce the negative charge and when you have a methylation that also is going to modulate the surface property of that particular protein. So as soon as you have the acetylation and phosphorylation you are going to have the unwinding of the chromatin structure and DNA becomes more accessible for the other kinds of downstream applications or other kinds of downstream activities such as replication, transcription and translations. So acetylations, so acetylation take place at the lysine residue of K4, K5 in the H4. It takes place through an enzyme which is called as histone acetyltransferase or AT. The acetylated chromatin are more open this means they are actually going to be active in terms of the replication and transcription.

It is accessible for transcription factor and polymerases. Deacetylation take place by the histone deacetylase or HDAC. The acetyl group donor is acetyl-CoA. So you are going

to have the closed chromatin and when you have the activity of the histone acetylase it is actually going to acetylate the chromatin and that is how it is actually going to form the relaxed chromatin. So once the relaxation is over when that process is over then you are actually going to put the HDAC activity which means the histone deacetylate and then it is actually going to be closed chromatin and that is how these are the things which are actually going to occur simultaneously or you know to just to you know unwind the DNA make it all accessible so that you can be able to use that DNA and then after that once that process is over then you can actually be able to close that DNA.

Same is true for the phosphorylation. So you are going to have the kinase activity which is actually going to convert the closed chromatin into the relaxed chromatin and then you have the phosphatases which are actually going to remove the phosphorylations and that is how it is actually going to reverse the events. Then third is methylation. So it occurs on the side chain of the lysine and arginine. The methylation does not alter the charge but it actually changes the charge what is present on to that particular residues. So lysine can be mono methylated di or tri methylated methylation done by the histone lysine methyl transferase and the histone lysine methyl D methylase.

So you are here also you are going to have the methyl transferase and the D methylase and when you have the methyl transferase the closed chromatin is going to be converted into the relaxed chromatin and same is going to be reversed by the histone lysine D methylase. So these are the about the normal chromosome. These are the information which is required for the normal chromosomes but when people were discussing about or when we started investigating the different types of chromosomes they could find some are the specialized chromosome which is present in the some of the organisms. So let us discuss about this specialized chromosomes and how the DNA is packed into the specialized chromosome and what are the different properties of these specialized chromosomes.

So the first is polytene chromosomes. So what are polytene chromosomes? Polytene chromosomes which are gigantic chromosome that grew from a smaller developing chromosomes frequently appear into the slavery gland of the pteran flies such as *Drosophila melirulogaster*. They are also known as slavery gland chromosome because they were found in slavery glands. The Balbini found a polytene chromosome in the slavery gland nuclei of the larva in the 1881 due to the presence of several chromatin in them they are known as the polytene chromosomes. Now the question is how these polytene chromosomes are being found? Though the most polytene chromosomes are located in the interface nucleus of a few cell in the pteran fly larva each chromosome component is successfully duplicated as they grow from the chromosome of the duplicated nucleus. After each DNA doubling the later stage of the mitosis are

eliminated leading to the deployment of a polytine chromosome as a result the cell cycle is divided into the S phase and G phase in *Drosophila melirulogaster*.

This polytineation cell cycle is developed during the mit embryogenesis. DNA strands do not separate at the final stage of each S phase rather they remain accompanied to one another generating the polytine chromosomes. The process of endor duplications or multiple chromosome DNA replication without adequate kynokinases and cytokinesis results into the polytine of gigantic chromosomes as a result the giant chromosomes are produced which are 70 to 100 times longer than the typical metasource chromosomes. Morphological features of the polytine chromosomes. The polytine chromosome is a very important thing actually because this is a kind of an exception or kind of the structure what is been found in a specific organisms.

So there are numerous partially duplicated chromosomes that are almost intervened with each another making up the polytine chromosomes. The heterochromatized centromere of all chromosome fuse in a centromere. The polytine chromosomes are found in the form of 6 radiating arms from the chromo center you can have the X chromosome you can have the 2 chromosomes, 2 chromosomes left arm right arm, third chromosome right arm, left arm then you can have the fourth chromosome which is the shortest arm then you can have the Y chromosomes and so on. An altered pattern of bright and dark is seen when these chromosomes are stained and examined under a microscope. Inter-band referred to a light pattern which band referred to the dark pattern.

So this is a specific polytine chromosomes where you are going to have the right arm you are going to have left arm within the left arm you are going to have the left arm of chromosome 2, left arm of chromosome 3 similarly you can have the right chromosome of arm 3 and so on. So within this place your centromere you are going to have the divergence and that is how you are going to have the X chromosome and Y chromosomes and so on. Then you can have the some of the classical characteristics of these chromosomes such as you are going to have the Balbini rings to the band undergoes morphological and biochemical changes related to their gene activity and the activation of the genes of a band causes the compact chromatin strands to uncoil and expand outward resulting in a chromosomal puff. The puff contains the DNA loops that are less condensed and the DNA of band elsewhere in the chromosome puffs are active gene of the transcription. So these are the puffs so chromosomal puff and these are the active region of the gene expressions.

What is the function of the polytine chromosomes? The nuclei of each cell enlarge in size leading to a cell growth. The metabolic benefit of having a numerous copy of a gene allow for a higher expression of gene expressions. The chromosome in *Drosophila*

menagaster undergoes numerous round of endoreduplications in order to generate the significant amount of glue prior to pupillations. The bar phenotype which includes the kidney shape eyes occurs from the tandem duplication of the severe polytine bands that are close to the centromere of the X chromosomes due to the fact that the polytine chromosomes are interface chromosomes and are thus transcribed. As a result it offers a chance to investigate transcription by the direct observation and transcriptional response to the certain stimuli can be observed.

So apart from the palatine chromosomes you can also have the another kind of chromosome which is called as Lambrech chromosomes. So Lambrech chromosomes what are Lambrech chromosomes? Lambrech chromosomes are transcriptionally active chromosome which are mainly found in the germinal vesicle of large courses of many vertebrate and the invertibrate. The Lambrech chromosomes derive their name from the lateral loop that exclude from the chromosome at a certain point. They are very much transcriptionally active as the emerging DNA from the certain point are rich in RNA polymerases. These chromosomes were first observed by the Fleming and Ruckerts in 1882 in oocyst of the amphibians.

Where do these chromosomes occur? The Lambrech chromosomes occur in the diplodin stage of the prophase I of the first meotic division in the primary oocyst of all the animals and the structure of the Lambrech chromosomes. So each RNA polymerase is attached to the nascent RNA and associated protein generating the visible brush like appearance. It can be visualized easily that Lambrech chromosomes are held in a stretched out form during the diplodin stage of the prophase I of the first meotic division. The axis of Lambrech chromosome contain array of beat from which the loops are protruding onward called chromosomes. They exist as meotic bivalent homologous chromosomes held together by the Crease meta.

So these are the so in this Lambrech chromosomes you are going to have the chromosomes then you are going to have the RNA polymerase which is protruding towards. So this is actually going to be a region of singly chromosomes and so on. And it is actually going to be transcriptionally very active because these are the region which are actually going to be available for transcription and the translation. So they contain the symmetrical loops one each other chromatin in a chromosome their absence of lesion in the centromere region each loop and each loop bears an axis which made up of single DNA molecule that is unfold during the RNA synthesis. What is the function of the Lambrech chromosomes? Prokaryotes are useful in chromosomal mapping then extremely helpful in visualization of gene expression and also the change associated with the transcription.

It provides a great proof for the eukaryotic gene amplification which play a crucial role in the oocyte development and it is helpful in the hybridization results. Now at the end we are going to discuss the comparison of the prokaryotic as well as the eukaryotic genome. Many of these properties we have already discussed right. So the comparison of the prokaryotic as well as the eukaryotic genome. So prokaryotic genome is small in size it is going to be large in the case of eukaryotic genome.

Genome is going to be a DNA and a few protein in a simpler manner whereas, DNA in the case of eukaryotic cell the genome is going to be present and many proteins are involved such as the histone proteins and so on. It contains a single set of chromosome whereas, in the eukaryotes you can have the multiple set of chromosomes. The amount of DNA is going to be small in the case of prokaryotic genome it is going to be typically very large number of DNA which is present. Then the prokaryotic genome is polycystronic whereas, in the case of eukaryotic genome it is monocystronic. Then most of the DNA encodes for the protein so it is actually there is no useless DNA right there is no DNA which is not going to be transcribed or which is not going to be translated to the protein whereas, most of the DNA does not code for the protein right.

It is very a small portion of the genome which is actually coding for the protein rest all is non-coding regions. Then RNA processing not an option right so RNA processing allows for the several of these genome because you have the non-coding regions is non-coding region has to be separated from the coding region and that is why the RNA processing is required. And messenger RNA has a short life span whereas, the messenger RNA is long life because the eukaryotic cell requires the continuous synthesis of a protein for several days right. So these are the some of the properties of the genome we have discussed about the prokaryotic genome and the organization of the prokaryotic genome and we have also discussed about the eukaryotic genome and eukaryotic genome organizations. We have discussed about the organization of the proteins of the different protein which are involved in the nucleosomal assembly and how the nucleosomal assembly is being formed and so on.

So with this brief discussion about the genetic material we are going to conclude our lecture here in our subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 04
Central Dogma of Molecular Biology
Lecture-20 Central Dogma of Molecular Biology

Hello everyone, this is Dr. Vishal Trivedi from Department of biosciences and bioengineering IIT Guwahati. Once we have understood the biomolecules and their structure and functions, it is important for us to understand the how these biomolecules are playing the role in different types of activities. So, let us take an example, right, if I think we have discussed this kind of activity when we were discussing about the carbohydrate metabolism. So, cell is continuously requiring the energy production and that energy production is being done by the burning the food reserves, right. So, one of the major food reserve which can be burned for running the metabolism is the carbohydrates.

And you might have seen that when we were discussing about the glycolysis or the grape cycle, we have discussed about the participation of the different types of biomolecules. Like one is bio, one is carbohydrate which is involved in the going through the different reaction intermediates. And to perform these reactions, you are actually having the coordinated actions of the different types of enzymes. And then apart from that, if there will be a shortage of any intermediate or there will be an access of intermediates, these intermediates can be pulled back to the different biomolecules.

For example, if there will be access of acetyl CoA, then it is actually going to synthesize the different types of lipids exactly. On the other hand, if there will be a shortage of acetyl CoA, then it is actually going to be provided after the beta oxidation and then it is actually going to channelize into the grape cycle. So, there is a coordination among the different biomolecules. And I think if you recall, we have also discussed about the different types of activities, what are happening when you are actually running the life cycle of an intermediate, right. So if you see the life activities, what you see here is that organism is considered to be life, if it actually performing the these four functions, it has the ability of self-growth or the self-renewal, it should have an ability of indigenous ability to produce the energy, then it should also have a movement, right, should have a movement from one place to another place.

In this case, there is an exception that the plants are not moving from one place to another place, otherwise, there are plants also which are moving from one place to another place. And then it should also have the ability to self-renew. Now if you see

how these events are actually happening, right. For example, if you see the self-renewal, right. So if we take an example of the bacteria, right.

And I think if you remember, we were discussing when we were discussing about the cell division, we discussed how the bacteria is going to replicate and going to provide the more number of bacteria. So if there is a single bacteria, what you have in the bacterial cell is actually a chromosome which is, so it is going to have the cell wall, right. And then it also going to have the chromosome, right. So it is going to have a nuclear content. And if this bacteria has to replicate and give you the two individual bacteria, then it first has to grow in size.

So it is going to grow in size, which means it is actually going to increase the amount of cytosol and then it also going to increase the amount of nuclear content. So what you see here is the nuclear content is going to be double, right. So this is going to be 2x nuclear content, right. And it is also going to have the double the amount of cytosol and then eventually it is actually going to divide into two, right. And it is actually going to give you the two different bacteria.

So now, what you see here is that it is actually going to have the action of the different types of molecules, right. When it wants to grow from one bacteria to a slightly larger in size, it is actually required to synthesize the different types of biomolecules. It requires the synthesis of the proteins. It also requires the synthesis of lipids because it has to synthesize the plasma membrane. And then it also required to synthesize the nuclear content, right.

So it is going to synthesize the nucleus, right. It is going to synthesize the DNA and as well as the RNA. And what you see here is that the synthesis of these biomolecules are going to be essential for this particular bacteria even to go for a simple process of division, right, where it is actually going to give you the two different bacteria. Similarly, we have discussed about how the endogenous ability to produce the energy. So if an organism is having an endogenous ability to produce the energy, it means that the organism is actually going to, you know, have the system, right, has to have a machinery, right.

So in this case, you can have the two different types of example, right. You can have the plants, which are actually going to perform the photosynthesis. And if you remember, when we were discussing about the chloroplast, we have discussed about the complete detail of the photosynthesis where we have the dark reactions and the light reactions. And in both the cases, whether it is a dark reaction or the light reactions, it is actually going to utilize and it is actually going to synthesize the different types of

biomolecules. It is also going to synthesize the different types of proteins, right.

It is requiring, you know, for light, for example, for light reaction, it requires the ATP synthase to produce the amount of ATP. And then for the dark reactions, it requires the different types of enzymes like rubisco and all other enzymes. So even if you see the photosynthesis, for photosynthesis also the plant has to synthesize the proteins. And it also requires the different types of receptors and channels to receive the atmospheric oxygens and then also requires the stomata, right. So it is opening and closing of the stomata is also going to be governed by the different types of biomolecules.

As far as the animal is concerned, right, animal, for example, if the animal is concerned, animal is also going to utilize the preformed food, right. For example, if you talk about the humans, they are actually going to utilize a very well developed system where they are actually going to digest the food material and into the elementary canal and then they are actually going to have the constituent molecules and then these constituent molecules are going to be channelized into the different types of metabolic pathways. We have also discussed about the glycolysis and Krebs cycles and how the different types of enzyme, lipids and you know, the carbohydrates are participating into the energy productions. And then it also has the ability to self replicate. So self replication is also where it is actually going to produce from the one person to, so one organism is going to produce the two different organisms.

So these are the few of the classical and the basic pathways which actually are essential and what you see here is that the synthesis of the new biomolecule or the synthesis of the different types of protein, lipids and nuclear content is essential for performing these functions. Apart from these functions, we have several more examples where also the life activities are involving the different types of production of the different types of biomolecule. For example, the adaptations. Adaptation is a very, very important phenomena for the survival of that particular organism. Remember that when we were discussing about the evolution, we discussed the different types of theories and whether you go with the Lammock's theory or whether you go with the Darwin's theory, both of these theories were heavily been dependent on the adaptations.

Like the process by which they have been explained the adaptation is different, but the adaptation is a very, very successful or where adaptation is an essential phenomena for an organism to succeed in their life cycle or to complete their life cycle so that they can be able to produce their offspring. So that adaptation is a multidimensional event. The second option is where you are actually going to get protection from the prey. So you can get the protection from the prey. You might have seen the example of the deer, how the deer has developed the strong muscles.

And because of that, the deer can be able to run very fast. And it happens because the deer has to get protection or get has to be remain safe from the tigers and other carnivores. So that is a classical example where the deer has developed the muscle cells. So if it is developed the muscle cells, this means it actually has produced the large amount of muscular proteins, the proteins which are responsible for making the muscle cells. And these proteins are actually being responsible for providing the extra power into the muscles of the deer and that's how it is actually getting protection.

Similarly, we can also have the different types of, for example, the adaptation also needs that you can have some kind of phenotypic changes. Like for example, if you can have the color of the skin change, so if there will be a skin color change, that may actually also going to do the similar kind of function that can also be able to protect the organism from the prey. And then it also can actually allow the organisms to go for the better choice for the sexual partner. So that is also very important because when you are actually going to have the better coloring patterns and when you have a better way of attracting your sexual partner, you can actually have the higher chances of going for the sexual activities and that's how you can be able to have the higher chances of producing the more number of offspring. And this also is requiring the production of the different types of biomolecules.

So different production of the molecules or biomolecules and eventually all these biomolecules are end up in one molecule which is the protein and the protein is because the protein is responsible for the providing the colors to the skin protein is responsible for making a different types of metabolites so that they can be able to give the patterns into the skin color and so on. So eventually what you see here is that whether we are actually going with the basic activities like the sexual reproductions or the running its own metabolism and all those kind of things or whether we are going with the specific activities where you are actually requiring the adaptations or phenotypic changes or running its own your own metabolism, the ends comes when there you are actually going to produce the different types of proteins and the protein production is directly linked to the generation of the amino acid. Which means if you have to synthesize a protein you also have to synthesize the amino acids and then you have to connect these amino acid in a specific order. Remember when we were discussing about the proteins we said that the protein is a polymer of the amino acids where the amino acids are combined to each other by a peptide bond and these amino acid has to be put in a systematic or in a sequence which is actually going to be defined the secondary structures and then the secondary structures are going to define the tertiary structures and the tertiary structures is eventually going to define the function of that particular protein. So what you see here is that if a protein production is actually being essential then the protein is going to be

provided or thus going to be synthesized by synthesizing the different types of amino acids and then combining them in a particular sequence and that is being done by a systematic series of the reactions.

And all these reactions are coming under the big umbrella of the central dogma of life. So what is central dogma of life? So what central dogma of life is that the protein is made up of the amino acid and every protein has the unique amino acid arrangement in a specific sequence. But we have seen the many examples where the protein is important for the particular activity of that particular organism. The information to synthesize the protein with a unique amino acid sequence is provided by the nucleic acid present within the nucleus. In a preset sequence the DNA present in the nucleus give rise to the specific RNA sequence and in that in turn guide the cellular machinery to synthesize the protein.

So what will happen is that the DNA what is present into the nucleus or whether it is present into the cytosol like in the case of prokaryotic cell it is actually going to synthesize the RNA and then this RNA is actually going to have the sequence information to synthesize the different types of amino acids and then these amino acids are going to combine to give you the specific proteins. That is why if you see if you want to have the proteins you first have to synthesize the RNA and then you have to synthesize the DNA which means it has to first synthesize the RNA from the DNA and then from the RNA it is going to synthesize the DNA. Scientists consider this as the fundamental event to run the life and considered as the central dogma of life which means it is actually going to be the central theme of any kind of life related activities. We have seen the many types of life related activities and where the everywhere you have to have some kind of production of the protein and if you want to see the production of the protein you might have to synthesize the RNA and that RNA has to be synthesized from the DNA. In another word the Francis communicated to the journal Nature state that the central dogma of the molecular biology deals with the detail residue by residue transfer of sequential information states that the such information cannot be transferred back from the protein to either protein or to the nucleic acid which means that the central dogma of life is also been called also been known as the central dogma of molecular biology and that is going to as well the you know the statement given by the Francis it says that it is actually going to deal with the residue by residue transfer of the sequential information which means you are going to have a residue of DNA which are going to synthesize the RNA and then it is also going to synthesize the protein.

What he also says is that the reversal is not possible which means from the protein you cannot synthesize the RNA and from the RNA you cannot synthesize the DNA which is we know now that it is not true right in the under the special circumstances you can be able to synthesize from the RNA to DNA and as well as from the protein to RNA as well

or from the protein to DNA as well that you are going to see when we are going to discuss more about the central dogma of life. So, what you are going to see is that if we have to follow the central dogma of life we have to run the multiple reactions. So, the central dogma of life is the basis of life on the earth and it is required to control the biological processes. Following this, hypergial flow of information from the DNA to protein allow the nucleus to control the all biological activity in a cell. The normal condition the flow of information sequence to sequence requires the three processes.

Now the question is why there is a requirement of the central dogma of life or the central dogma of molecular biology. It is required because then you can have the full control over the all the activities of a particular cell because you cannot have the that particular type of biological process started or ending until you have the particular protein to be produced. Like for example, if we want to start the glycolysis, we have to have a synthesis of hexokinase or the glucokinase, but you cannot have the synthesis of the hexokinase or the glucokinase spontaneously. It has to be governed from the DNA and that is why it says that if since everything will depend on the DNA to first synthesize the RNA and then the RNA is going to synthesize the DNA, the RNA is going to synthesize the protein, it is actually going to allow the DNA what is sitting inside the nucleus to control all these processes and because of that the nucleus can be able to control all the biological processes. The three processes which are required for the flow of information is as follows.

You can have the sequence dependent synthesis of DNA from the pre-existing DNA. So you can have a DNA and that DNA also has to be synthesized. Remember that mean there will be a duplication of organism, the organism want to go for the division whether it is a bacterial cell or the eukaryotic cell, it has to have the duplication of the DNA. So you can have the DNA dependent DNA synthesis and that is actually going to be performed by an enzyme which is called as the DNA polymerase and the process which is by which you are going to do this is called as the replications. Then you can have the sequence dependent synthesis of RNA from the DNA.

So once the DNA is synthesized then you can also have the DNA dependent RNA synthesis and that is also going to perform by the enzyme which is called as the DNA polymerase, RNA polymerase. So once the RNA is and this process is called as the transcription. Then we have the sequence dependent synthesis of DNA from the pre-existing DNA. So once you have the RNA then RNA is going to be have the synthesis of the protein from the RNA molecule and that process is called as the translation. And this process is called as the translation.

So if you want to run the central dogma of life or if you want to continue the all the

biological activities within the cell, you have to follow these three events. You have to synthesize the DNA from the existing pre-existing DNA and that process is called as the replications. Then you can have the sequence dependent synthesis of RNA from the DNA that process is called as the transcription and then you can also have the sequence dependent synthesis of DNA from the pre-existing DNA, sequence dependent synthesis of the proteins from the pre-existing RNA and that process is called as the translation. Now we can have the more detail about these processes. So replications, so genomic content in an organism need to be duplicated from the S phase of the cell cycle.

Transcription of DNA is done by the replication utilizing the sequence information of the parent DNA. The enzyme used in this process is called as the DNA dependent DNA polymerase. Then we can have the transcription. So the DNA is present in the nucleus whereas the protein synthesis machinery is present in the cytosol. Whereas in the case of prokaryote the both the transcription and the translational machinery is both present inside the cell so because there is no nucleus.

Hence the information present in DNA is used to synthesize the RNA which has the ability to transport outside the nucleus to participate into the protein synthesis. Synthesis of RNA from DNA is done by the transcription utilizing the sequence information of the DNA. The enzyme used in this process is called as the DNA dependent RNA polymerase. Then we can also have the translation. The RNA present in the cytosol is utilized by the translational machinery to synthesize the protein in a sequence dependent manner through a process known as the translations.

So these are the things which we have depicted here. The DNA dependent DNA polymerase synthesis which is going to be done by the DNA polymerase. The process is called as the replication. Then you can have the transcription and then you also have the translations. But as I said you know although the Francis has said that these processes cannot be reversed or the sequence of information can be only from the DNA to RNA and then RNA to protein it cannot be reversed.

But under the specific conditions biological system does not follow the usual pathway to replay the information which means it also can have the reverse directions. So under the normal circumstances what you have is you have the DNA dependent DNA synthesis. This is called as the replication. Then you can have the DNA dependent RNA synthesis.

This is called as the transcription. And then you can also have the RNA dependent protein synthesis which is also called as the translation. So these are the normal circumstances which are happening in every cell. But there are an exceptional or special cases where you can have the other activities. For example the RNA dependent DNA

synthesis which is called as the reverse transcription.

So this is actually a reversal of this. This is reversal of this. So it is also called as the RNA dependent DNA synthesis. So that is called as the reverse transcriptions. In most of the organism the genomic content is present in the form of DNA. But in few organisms such as viruses RNA is also present as the genomic content.

And in these cases RNA need to be converted into DNA and replicate using the host machinery cycles. It is done by reverse transcription utilizing the sequence information of the parent DNA. The enzyme used in this purpose is called as the RNA dependent DNA polymerase or it is also called as the reverse transcriptase. So this is what is shown here. You can have the normal circumstances where the DNA is actually going to be synthesized by the DNA and is going to give you the by the replications.

Then the DNA is going to give you the RNA by the process of transcription. And then the RNA is going to give you the protein which is called as the translations. Or you can have the special circumstances where the RNA is actually going to give you the DNA and that process is called as a reverse transcription or the and the enzyme is called as the reverse transcriptase. Then we have another example where you can have the RNA dependent RNA synthesis and that is called as the RNA replications. In most of the organism the genomic content is present in the form of DNA but in some organisms such as viruses RNA is present as the genomic content and done by the replication utilizing the sequence information of the parent RNA.

The enzyme used for this purpose is called as the RNA dependent RNA polymerase. And then you can also have the DNA directly synthesizing the protein. So there will be a complete absence of this particular RNA species. So you can also have the DNA dependent protein synthesis. So in that case the DNA directly giving you the protein under the in vitro cell free system.

So it is not possible in the in the in an organism but it is also possible under the in vitro cell free system. So in a cell free system it is possible to translate the DNA directly into the protein in the presence of ribosome. It is not precisely controlled and it is not known whether it is synthesized the protein in a sequence dependent manner which means this is possible but it is not precisely controlled. It is not precisely going to give you the specific protein molecules.

It may actually give you the protein molecule. It may give you a random amino acids combined with each other. Then we also have another example where we can have the additional possibilities. So there we have an additional possibility like the protein is

giving rise to DNA or protein is giving rise to RNA or protein is giving rise to protein without going to the normal second chances of the central term of life. So we have a protein to protein.

These are considered as the infectious protein. They are also called as the prions. So in a prions what you have is that the protein is replicating from the protein itself. So in this these are considered as the infectious protein which replicates to form the identical copies on themselves. These proteins are known as the prions although it is represent the transfer of information but it is not the usual pathway and considered to be exception of the central dogma of life. So what we have discussed we have discussed about the importance of the central dogma of life or the central dogma of molecular biology and how and why there is a requirement of such a pathway.

Do we require that pathway so that the DNA can be able to precisely be able to control the production of RNA and in turn it also can actually have the full control over the protein molecule. Not only that the DNA is actually going to have the information which is going to be more stable. However that mean we are discussing about the DNA and RNA and as well as the protein what we said is that the DNA is a very stable molecule compared to the RNA because the DNA is having the double helix structures. DNA is having the deoxyribose sugar and it also has the more stability and that and because it is protected inside the nucleus and it is protected by the you know the some of the binding proteins it is very very preserved. So that is why the information what you have in the DNA is not having the not going to give you the any kind of alterations whereas the information what is present into the RNA is very susceptible for the any kind of damages and same is true for the protein.

The protein also can get modified and it can also get altered and all that. And on the other hand the RNA is very very you know not stable right RNA is less stable compared to the DNA because RNA is single standard, RNA is susceptible for the different types of RNases what are present in the cytosol and so on. So because of that the biology or the life has decided that okay I will going to utilize the information what is given into the DNA. But since the DNA cannot participate directly and give you the protein because if that is the case the DNA is actually not been able to synthesize the different types of proteins right because if suppose the DNA has to synthesize the thousandth of the proteins and then the DNA molecule is going to be remain open right and it has to keep synthesizing the protein. So because of that purpose the DNA has decided that okay I will take the help of the RNA.

So what it will doing is it is actually making its own copies in the form of RNA and that is how it is actually distributing the work right just like any remember that the if any for

example the prime minister of the country right prime minister of the country can is his have a is has the ability to perform the different types of task right. But it cannot do all those tasks on its own because it has a limited resources right. So because of that what it is doing is it is actually making the different types of portfolios and then he is distributing those portfolios to different types of ministers. So these are actually going to be the ministers for the DNA and that is how they are actually going to do their special function. For example if DNA has to synthesize the actin, myosin, beta, beta ATP synthase you know some described of different types of proteins it cannot do this function on a timely manner it cannot do that because it cannot channelize into the multiple types of activities.

So what it can do is it can synthesize the RNA which is for the actin molecule it can synthesize the RNA which is for the myosin it can synthesize the RNA for the ATP synthase and that is how these different RNA molecules are now going to be taken up by the protein synthesis machinery and that is how these RNA molecules are going to give you the corresponding proteins. On the other hand because the DNA is stable you cannot modulate the expression level of this particular protein right. But at the same time if the function of that particular process is over and suppose you want to degrade the myosin or ATP synthase then you cannot degrade that region of the DNA right. What you can do is you can just simply degrade that amount or that particular RNA. If you degrade the RNA you are eventually going to degrade the protein because the protein have a definite half life or definite age right after that their protein is going to be removed from the system right.

So that is also another aspect why the biology has decided to go for these you know these kind of sequential transfer of information from the DNA to RNA and RNA to protein. So central dogma of life is very important for us to understand the different types of activities what is happening into the cell. So this is all about the central dogma of life and what we have discussed we have discussed about the many events what is happening inside the cell and what is responsible and what are the different types of you know moly molecules are going to be produced. And what we have understood is that the protein is actually been responsible for making the different types of activities within the cell and the protein synthesis is always been governed by the sequence dependent synthesis of the RNA and the sequence dependent RNA synthesis is being done by the DNA. All these events are actually been coming together under the central dogma of life or to the central dogma of molecular biology as it is been stated by the Francis.

What we have discussed we have discussed about the applications like the DNA dependent DNA synthesis which is been done by the DNA polymerase and we have discussed about the different events where we have discussed about the origin of

applications. We have discussed about the initiation, elongation and terminations. So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss some more aspects related to the central dogma of life where we are going to discuss about the transcriptions. So with this I would like to conclude my lecture here. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 05
Replication

Lecture-21 Replication in Prokaryotic System (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati and what we were discussing we were discussing about the different aspects of molecular biology in this particular MOOCs course. So, so far what we have discussed we have discussed about the cellular structures, we have discussed about the structure and functions of the different types of biomolecules where we have discussed about the DNA, RNA, protein and enzyme and in the previous module we have also discussed about the central dogma of molecular biology. So, we have discussed that how the different types of cellular activities are dependent on the protein synthesis and how the protein synthesis is being regulated by the level of production of RNA and then by the production of RNA from the DNA. So, when we were discussing about the central dogma of molecular biology we discussed about that it is actually a regulated event of the three different types of activities where you are actually going to have the production of DNA from the pre-existing DNA and this process is known as replication whereas, the DNA is actually going to produce the RNA and this process is known as transcription and the third is the once the RNA is available then it is actually going to be translated into the protein and all these events are being catalyzed by the different types of enzymes.

So, replication is catalyzed by the DNA dependent DNA polymerase whereas, the transcription is catalyzed by the DNA dependent RNA polymerase and the translation is being done by the ribosomes. So, in today's lecture we are going to start discussing about the first event and then the first event is the synthesis of the new DNA from the pre-existing DNA. If you recall when we were discussing about the cell division we discussed that during the S phase the DNA is going to be synthesized and then you are going to have the two copies of the same DNA and then these two copies of the same DNA is actually going to be shared between the cells and that is how the cell is actually going to replicate and cell is going to divide. So, synthesis of the new DNA molecule from the pre-existing DNA molecule is known as replication.

Now, replication is an important event because it allows the synthesis of so replication means the synthesis of the DNA from the pre-existing DNA this means this is going to be this you can imagine that this could be a DNA of the parents and this could be a DNA of the offsprings. And you can understand that if there will be any this whole process has to be

done with lot of precision with lot of accuracy. So, that you are not going to carry the bad information right. So, you want to carry the same information what is present in this particular DNA and all the other hand it has to be complete in a given time period. So, replication has the two task one is the synthesis of the new DNA considering old DNA as template.

And the second is it should be accurate. So, there should be so it should be almost the exact 100 percent replica of the older DNA there should be no mutations or there be no substitutions there should be no alterations and all that. So, to achieve this the machinery has to be very robust machinery has to be very you know should have the components. So, that it should be able to do this job do the synthesis job, but at the other hand it should also have the components. So, that it should be able to do a quality checking and the replication is also related to the genome right.

So, prokaryotic genome and the as well as the eukaryotic genome if you recall in the previous lecture we have discussed about the prokaryotic genome and the eukaryotic genome and there is a significant difference between the prokaryotic genome and the eukaryotic genome right. And as a result the machinery is also different between the prokaryotic system and the eukaryotic system. So, in the prokaryotic system we have the different types of components which are required for synthesis of the genome whereas, in the eukaryotic systems because the eukaryotic genome is more complex you require the different types of machinery. So, because the machinery is different I have you know splitted this particular thing into two components. So, that you are we are going to discuss first about the prokaryotic replication and then we are actually going to discuss about the eukaryotic replication.

Because when we will discuss this in separate or when we are going to discuss about the prokaryotic versus eukaryotic you will understand that how the machinery is been adopted from the prokaryotic system to the eukaryotic system. So, that it is more efficient and it is actually bringing the more accuracy in the things. So, the first question comes what is replication? We already discussed right. So, DNA replication is a biological process that helps to transmit the DNA or the genetic information from the generation to generation by producing the two identical DNA strands of the daughter cells from the double standard parental DNA. This anyway we have discussed.

What is the importance of replications? Application is a way of duplicating all the genetic information from the all living parent organism to the daughter organisms thereby helping maintain the genetic materials intactness and the organism survival. Now, if this is such an important event and it may be depend on the type of DNA you would like to do the replications and so on. It is important to understand the genetic

material of the prokaryotic system because we are going to start the replication of the prokaryotic system first. So, prokaryotic genetic material remember that when we have discussed in detail about the prokaryotic replications prokaryotic genetic material. So, prokaryotic genetic material is a single chromosome and it is mean mostly been found in the cytoplasm right.

So, prokaryotic genome can be found in the cytoplasm. They are negatively super coiled and circular. Generally they can be found into a singular number. However, exceptions exist for example, the Vibrio cholera has the two copies of the genome right. So, two genomes prokaryotes such as E.

coli contains a self replicating extracellular genetic material which is also called as plasmid. So, you have the circular genome right or circular chromosome. So, it is called chromosome right. So, see you have a single copy of the chromosome which is negatively super coiled and it is a circular DNA. And mostly there are only one copy of that circular genome is present in the prokaryotic system, but there are exceptions that in the Vibrio cholera you have the two genomes.

Apart from this chromosome you are also going to have the extra chromosomal self replicating genome and that is called as plasmids. And when we talk about the genetic material we are actually going to talk about the submission of the chromosome and as well as the plasmids. Now, as far as the plasmids are concerned plasmids are of small size DNA it is only in the 1500 to 2000 kB. Therefore very small compared to the actual prokaryotic DNA they contains the origin of applications and it replicates independently. So, plasmid does not require any kind of help from the chromosome they are independent self replicating extra chromosomal genetic material.

Now let us first ask the questions how you can actually be able to have the replications right. So, there are three different modes which are being proposed as far as the replication is concerned. So, replication can be done in the three different mode one is called as the conservative mode, the second is called as the semi conservative mode and the third is called as the dispersive mode. So, in a conservative mode conservative mode produces that or if it is conservative mode then it produces that two DNA helices from the one single original DNA helix. One helix contains the downright prior ends DNA while the other contains the entirely new DNA.

Accepted postulate, but does not have that much significance. So, what is mean by the conservative mode is that from the parental DNA you are going to have after replications. So, if you are going to do the replications you are going to generate. So, for example, from the two strands you are going to have the four strands right. So, four

strands are actually going to be segregated and as a result you are going to have for example, here you have 1 and 2.

So, that 1 and 2 will go into the parents whereas, the newly synthesized 3 and 4 will go into the daughter. So, as this is called as the conservative mode. In a conservative mode there will be no mixing of the content from the parents DNA. So, there that is why in a second round of replications you are going to have the parent DNA and you are also going to have the daughter DNA separately. This means there will be no mixing of the previous copy and it is actually going to have the pure DNA present in the parent DNA.

Then the second mode is called as semi conservative. So, one of the two helices forms each containing the one new and the one parental strands according to the Watson and Crick one strand serves as a template during the replication mode. The newly accepted postulate as the DNA polymerase needs a strand to form a new complementary strand. So, what is mean by the semi conservative mode is that from the parent DNA for example, you have the parent DNA as 1 and 2 it is actually going to do the replications and once you do the replications it is actually going to form the four copies right 1, 2, 3 and 4. So, 1, 2 is the original copy 3 and 4 is actually going to be the newly synthesized copy because the one from the one you are going to have the copy number 3 from the two you are going to have the copy number 4.

Now in the semi conservative mode the one is actually going to make a pair with 4 and the two is actually going to make a pair with 3. So, this is actually going to be 1, 4 and 2, 3 right. This means there will be a dilution of the genetic material of the parents and that is how the parents are actually going to share the 50 percent copy with the offsprings. That means, the two offsprings what are going to be produced after the first replication or the two DNA molecule what is or going to be formed after the first replications are actually going to have the mixture. It means going to have the 50 percent from the parents and the 50 percent of the newly synthesized DNA.

And this continues because even if I have a second replication this continued right and you are going to have the that is kind of dilution right. So, in the first replications you are going to have the mixing of the DNA from the parent DNA. So, you are going to have the two strands of the parents which are going to be diluted. Now, the third method is called as the dispersive. So, this would generate the DNA helices with the alternating pattern of the old and the new DNA segments and it does not have any kind of biological significance.

So, what this mean is that it is in the dispersive what we it says is that you are going to have the parental DNA and after the replication 50 percent portion is going to be

replicated with the original DNA and the 50 percent portion is going to be replicated in the with the new DNA. This means you are going to have a mixture of the both the strands and even the both components. This is not acceptable because it is not possible because when you have the copy number 1 and 2 you are going to have the generation of the 3 and 4. You are not going to have the 1 and followed by 3 and followed by 1 followed by 3 like that. So, that kind of scheme is not possible because once the DNA polymerase will sit and start synthesizing the DNA it will not come off and if it comes off then it is actually not going to have this particular type of pattern.

So, this is the most unacceptable method. Now, the question is we have the three different modes right. We have the conservative mode, we have the semi conservative mode and we have a dispersive mode and these are the three different types of model what has been proposed how the replication is going to happen right. But the question comes how scientifically we can be able to prove this. So, to prove this there is a simple experiment and the classical experiment what is being done by some of the scientists.

So, the Meselson star or Matthew Meselson and Stahl actually performed the experiments where they have asked what would be the mode of replication in the bacterial system. So, what was the aim? The aim was to establish which method is applied into the prokaryotic replication is it the conservative mode or semi conservative mode or the dispersive mode right. So, to ask these questions they have prepared the they have designed an experiment right and these are the requirements. So, you require a media you required a media to grow the E. coli for several generations and the media is having the two different types of media.

You are going to have the media with the standard form of nitrogen that is the N14. So, I will say this is the normal nitrogen that is the N14 nitrogen and then you also require a medium with the rare and the heavy form of nitrogen and that is called as N15. So, these are the two different types of isotopes what you can actually use. One is the normal normal nitrogen or the heavy nitrogen and then you can actually be able to you require the analytical reagents right. So, you can actually require the techniques so that you can be able to separate the DNA what is being formed by the N14 versus N15.

So, you require a cesium chloride gradient to get the separate DNA based on their density. Remember that this is heavy and this is light. So, it is actually going to have the different types of densities and that is why N14 can be separated from the N15 it is if it is present in the same mixture. So, if you have the pure DNA it is of the N14 or the pure DNA of N15 it is going to be formed the different bands into the cesium chloride gradient mixture. And if you are not sure about the gradient centrifugations remember

that we have already discussed that technique while we were discussing about the cell cellular fractionations and we have discussed about the density gradient centrifugations.

So, there we I think we have taken an example of the sucrose, but here we are they are using the cesium chloride. So, that is the only difference and you know that the gradient can be up in the upward direction or the downward directions. And we have also discussed in detail about the principles and how the things are getting separated when you are running them onto the gradient. So, gradient is actually going to separate the molecule based on their density. So, this is the procedure what you are going to follow right.

You are going to grow the culture of E. coli N15 isotope media for several generations. So, that the heavy isotope is incorporated into the purine and pyrimidine bases remember that both the most of the nitrogenous bases are made up of the N15. So, you when you grow them into N15 media all your purine and pyrimidines are actually going to be get labeled means all the nitrogen what is present as N14 is actually going to be replaced by the N15. So, all the DNA is going to be heavy DNA.

And then some of the E. coli from the heavy isotope media is taken and transferred into the normal media containing the N14 as the nitrogen source. So, now what will happen is that when it is going to synthesize the new DNA molecule it has to utilize this N14 right. So, based on the density gradient the DNA band will be generated into the centrifuge tubes. For equilibrium density gradient analysis the DNA are collected from the media and put it into the 6 molar cesium chloride. DNA sample in heavy salt gradient are taken into centrifuge for 50 to 60 hours at 1 lakh g rotations.

So, this is what exactly going to happen when you are actually going to have the E. coli into the N15 media you are going and when you run them onto the onto the cesium chloride density gradient what you will see is that it is actually going to form a separate band and this band is for the N15. So, yes when you are going to transfer that into the N14 you are going to have the bands which is of the intermediate density. So, this is the band this is the region where you are going to get the band for the N14 and this is the region for the N15. So, when you are going to grow the media when you are going to grow the bacteria into the normal media after N15 media it is actually going to have the N15 band into the intermediate spin right intermediates position it is not going to be related to the N15 or the N14.

Now if you grow the further right if you go for several more generations then what will happen is that you are actually going to start getting the band which is corresponding to the N14 which means there will be some DNA what is going to be served as the template

of the N14 and that is how you are going to have the intermediate DNA, but you also going to have the N14. Now this will continue and that is how this is going to be the final product. So, if you are only growing the bacteria into the heavy media you are going to have the single band at the N15, but you if you are running it at the into the N14 media. So, the DNA is initially been N15. So, N15 DNA is been transferred into the N15 labeled DNA containing bacteria when you put it into the normal media it is going to have the 50 percent N14 and 50 percent N15 and as a result after the first generation you are going to see a band which is intermediate So, this is actually N14 by N15 media right and this is going to be the intermediate DNA.

Now if you continue the for the several more generations then it is actually going to form the N14 dimers and as well as the N14 and N15 intermediate. So, it is actually going to form the intermediate DNA band and as well as the N14 light bands and this will continue because this is the actual thing right. So, this remember that when we were talking about the semi conservative mode what we said is that if you have this DNA right and you have the two strands at one and two and after the replication what will happen is that it is actually going to form the four strand right it is going to form two strands and it is going to form the four strands right. So, one is going to make a pair with four and two is going to make a pair with three this means the imagine that this is N15. So, if this is N15 so this is going to be N15 this is going to be N15 right.

Now here the new DNA what is going to be formed is not going to be N15 because there is no N15 available into the media right. So, what will happen is that the N1 strand which is going to be N15 the other strand is going to be N14 right and same is true here right N15 and N14 and that is why it is actually going to give you a intermediate density. Now if this will go for another generation then this N14 is going to replicate this N14 is going to replicate and N15 will again going to have another copy. So, this is actually what happen is that when you are going to have another replication then N14 and N14 and you are going to have the four DNA molecule right.

So, you are going to have the N15 as an N14 right. So, two copies of this and you are going to have the two copies of this right. So, he this is actually going to give you a band which is corresponding to the N14 and this is your this is going to give you a band which is for the intermediate. So, that is all you are going to get this actually here in the subsequent generations. So, what you can conclude from this is that whenever you there will be a DNA applications going to be impure it is going to give you a DNA which is impure. So, it is going to have the original copy and it is also going to have the new copy.

So, out of these three proposed replication mode the semi conservative mode is the

selected mode that can be observed into the prokaryotic system and as well as the eukaryotic system. Now let us talk about the replication machinery. So, first is what is replicon? So, replicon is the region of the DNA what is going to be participate into the replication reactions. So, replication replicon or the small stretch of DNA which is going to be involved into the replication is a DNA segment of the spore period that undergo replications. Replicants what is there in the replicon? So, replicon is going to have the origin of the application and it also going to have the terminations regions.

For example, the E. coli has one replicon on its genome. So, it is going to be mono replinonic which means it is going to start from one side and it is going to end on to the one side. So, it is going to have the mono replinonic whereas, eukaryotes have the multiple origin of replication hence they are multi replinonic. So, this means if it is equal if it is a eukaryotic region eukaryotic genome then you may have the multiple origin of replications and why it is so, because the eukaryotic genomes are big compared to the prokaryotic genome. So, they can actually afford to grow with go with the single origin of replication.

So, that by the time the replication is going to be over they are also going to have the synthesis and the other kinds of preparations. So, let us first talk about the origin of replications. So, origin of replications as the name suggests one the origin of replication represent the starting point of the replication in the prokaryotes. It is approximately 245 base pair 80 rich regions cis acting sequences. What is mean by the six acting sequences? Six acting sequences is that they can affect only the molecule of the DNA in which they reside.

So, cis acting means they are actually going to affect within that molecule. Second question is why it is 80 rich? So, 80 rich sequence is preferred because 80 rich is easy to melt. We know that A is making 2 base pair with T whereas, G is making 3 base pair with C. So, if you have the G into GC regions then it is you are supposed to break the 3 hydrogen bonding and it is difficult to break the 3 hydrogen bonding compared to the 2 hydrogen bonding. So, for melting or unwinding of duplex DNA less energy is required to break the hydrogen bonding than the GC sequences.

Melting of duplex DNA is ATP hydrolysis dependent as the energy released helps break the hydrogen bonding between A and T. Above mentioned 80 rich sequence is recognized by the enzyme which is called as DNA helicase to initiate the unwinding process. So, this is the typical original applications where you are going to have the 200 base pair 245 base pair long stretch and it is going to have the 80 rich region in the in the region and it is going to have the DUE and DOR. DUE means the duplex unwinding element whereas, DOR is called as DNA oligomerization.

So, the origin of the application in E. coli is known as the ODC. It contains the 2 short repeat motifs like 5 copies of the 9 mer sequences spread throughout the origin of the application DOR site serve as the DNA binding site of the DNA A which is a replication initiation protein and the 3 copies of the 13 base pair or 13 mer 80 rich repeat which is called as DO site where the DNA starts unwinding. So, upon binding of DNA A at the 9 mer region the 13 mer region starts melting. So, origins ORE of C contains the 11 copies of 5 prime GAT C 3 prime repeat methylated on that on both strands and only complete methylation can lead to the initiation of applications. Hemimethylated origins cannot initiate the replication until it is fully stored or methylated.

So, this is important that the methylation is also going to control the replications because remember that the methylation is the defense mechanism. We are going to discuss about some of these enzymes. So, which are I think going to be used in the molecular cloning. So, restriction enzyme restriction methylate system is a kind of a defense system and that is how the machinery is not going to initiate the replication until the DNA is fully methylated. If the DNA is not methylated, hemimethylated or unmethylated then it is not going to be replicated because it is considered to be host it is going to be considered to be DNA of the external origins.

Then we have the replication forks. Replication fork is going to be formed and that is how it is actually going to initiate. So, once the DNA is been melted at the origin of replication it is going to form the replication fork. So, the Y shape structure is generally found when the DNA starts melting and opening up. As the DNA open up bi-directionally at the origin of C2 replication forks are generated. Extension of the two oppositely directed replication fork leads to a replication bubble.

So, when the replication fork moves in this direction and as well as in this direction it opens the DNA in both the direction and as a result it is actually going to form a bubble like situation. Replication fork is going to move in both this direction and this is going to be called as replication bubble. And this replication bubble will move in both the direction where it is actually going to start synthesizing the new DNA. So, this is going to serve as a template for the this machinery and this is going to be served as the machinery for the this replication. So, you are going to have the one machinery which will run in this direction and they were going to have another machinery which will run in this direction for and that for it is actually going to have the replication of the both the origins, both the sites.

So you are going to have the replication in this direction and you are going to have the replication in this direction. To maintain the single standard situation the single standard

DNA binding proteins are going to code the single standard DNA to prevent the rewinding of the double standard DNA. So, these are some of the components what are going to be present within the DNA structure. So, you are going to have the replication original replications and then you it is going to form the replication forks. Apart from this you also require a battery of the machinery what is been formed or what is been assembled on to the replication fork to start the replications.

So, these are the some of the important enzyme what is been found into the prokaryotic replications. So, first enzyme is the DNA helicase or helicase right. So, it is called as DNA B it melts or open up the DNA at the replication forks. So, it is going to be the first enzyme which is going to be sit at the original replications and then it is actually going to start the opening of the DNA. Then the second is the single standard DNA binding proteins and these are they will prevent the unwinding of the single standard DNA to the double standard DNA.

So, you can imagine that as soon as the bubble is going to be formed the single standard DNA binding protein will go and sit on to the nucleotides. So, that they should not have any kind of interaction remember that these two molecules are complementary to each other. So, as soon as they get open they are supposed to be remain like that. So, that the other molecules will come and sit and do their jobs, but if you do not do that then they will come and stick to each other because your every A is going to have a you know complementity to T and every G is going to have the complementity to C. So, this complementity can only be break if you have a molecule which is sitting on to this right.

So, that is how you are going to have no interaction between these two and that is how they are going to be remain separated and then the other machine G will come and sit and start the replications. Then you also have the topoisomerase or DNA G. Topoisomerase works at the region ahead of the replication fork to prevent the super coiling. And then you also have the DNA pol I or it is also called as the Kornberg enzyme because the Kornberg is the first scientist who discovered the DNA pol I.

So, the first DNA polymerase to be discovered in E. coli by the Nobel prize winner Arthur Kornberg and the gene what is actually coding that DNA pol I is called as pol A and it is a monomeric protein of 928 amino acid or the 109 kilo Dalton and it has three different types of activity. It has a 5 prime to 3 prime polymerase activity, it has a 3 prime to 5 prime exonuclease activity and it also has the 5 prime to 3 prime exonuclease activity. So, do not worry about these activities because that anyway we are going to elaborate or going to discuss when we are going to talk about the other events. So, these three activities are really important. This activity is important for the DNA synthesis and other two activities are required for the proofreading and as well as the other kinds of

activities.

DNA polymerase I or the Kornberg enzyme is having the 5 prime to 3 prime exonuclease activity is independent of the other activities. Then we have the protease cleaves polymerase I into the two fragments right. So, if you treat the DNA pol I with the protease like trypsin then it is actually going to generate the two distinct fragments. And you are going to have the larger C-terminal fragment or it is called as Clino fragments which contains both the 5 prime to 3 prime polymerase activity and 3 prime to 5 prime exonuclease activity. And then you also going to have the smaller fragment which has the proofread active proofreading activity that is the 5 prime to 3 prime exonuclease activity.

So, Clino fragment is also very important and very popular in terms of the in vitro replication such as the PCR. So, sometime people are using the Clino fragments rather than the complete enzyme because the with the complete enzyme you have the always a danger of you having the 5 prime to 3 prime exonuclease activity and that may actually have the interference in terms of the DNA replication under the in vitro conditions. So, it has a low processivity which means only the 200 nucleotides can be processed then it also has the low polymerization rate. So, it has around 20 nucleotide can be added at the per second and the apart from these you also going to have other kinds of activities with associated with the DNA pol I that is the RNA primer removals, gap filling and the DNA repairs. These three are actually going to be discussed in detail when we are going to talk about the elongations and terminations.

And it is actually a metalloenzyme. So, it is actually a zinc dependent enzymes. Then we also have the DNA pol II. So, monomeric protein with pol B as a structural gene which is a size of the 90 kilo Dalton. DNA pol II having the two activities 5 prime to 3 prime polymerase activity and 3 prime to 5 prime exonuclease activity. It has a low polymerization that is the about 40 nucleotide per second and it also has the low processivity rate such as the 1500 nucleotides.

It mainly serve as the alternate DNA repair polymerase. Therefore it can be replicate DNA if the template is damaged. It does not require ATP for any type of this activity. Then you also have the DNA pol III. So, DNA pol III is a primary DNA replicase with the structural gene pol C and it has a 900 kilo Dalton. It is a multimeric protein complex of 10 different protein polypeptides such as alpha, epsilon, theta, z, tau, y, sigma and all that.

It has the high polymerization rate about 1000 nucleotides per second and it has high processivity that is the 50000 nucleotides. And DNA polymerase III serves as a

holoenzyme during their applications and holoenzyme refer to the multi protein complex whose catalytic activity is associated with the extra components. Then we have the DNA pol III. So, this is the structure of the DNA pol III and you can see that all these components are assembled to form this particular enzyme.

And there are four essential components. You have the two copies of the catalytic core that is the alpha subunit and the epsilon subunit. Alpha subunit has the 5 prime to 3 prime polymerase activity whereas the epsilon subunit has the 3 prime to 5 prime exonuclease activity whereas the theta subunit is for increasing the efficiency of the of the epsilon subunit. Two copies of the dimerization component that is the tau and the two copies of the homodimer of beta subunit ring for the processivity component and the one copy of the clamp loaders and these are the enzymes these are the subunits what are present in the clamp loaders. And how it is actually going to do the DNA polymerizations? So, the clamp loader links the two catalytic cores and the two beta clamps increases the processivity of the DNA pol III holoenzyme and the loading of clamps are done by the clamp loaders and the dimer component helps the catalytic cores to function at the same time.

Then we have the DNA polymerase IV and DNA polymerase V. Both are the Y family polymerase that do not have the 3 prime to 5 prime exonuclease activity. It has the low catalytic efficiency and a low processivity and fidelity. It is involved in the transition synthesis and replication damage DNA by the bypassing the nucleotide that can block the progression of a replication fork and it is been in size by the structural gene like the din B and the u mu d 2 c or the pol V. Then we also require the DNA primase. So DNA primase is required for the synthesis of RNA primers complementary to the strands then you also require the DNA ligase seal the DNA fragment gap into the strands.

Now there are 3 major events what is going to happen when you are going to have the DNA applications. So you have the first stage that is called as the initiation the second stage which is called as the elongation and the third stage is called as termination. So the first stage is going to have the recognition and starting of the replications then the elongation you are going to have the replication fork leads the dNTP synthesis and the proofreading and the third is termination where you are going to have the stopping of the replications. So E coli chromosome DNA is circular with no free end and it is replicated bi-directionally. So it resembles the Greek letter theta hence this replication mode is also known as the theta replication.

It can be seen in the gram negative bacteria such as proteobacteria some commonly used plasmid like Col-E1, RK2, F and P1 bacteriophage as well. So this mode of replication which is called as theta replication is been found into the prokaryotic system and as well

as the plasmids like Col-E1, RK2, F and P1 bacteriophage as well. So these are the some of the steps into the initiation. So initially the 2 to 4 DNA A protein using the ATP binds to the 9 mer DOR region in the ORAC performing the initiation complex. Once this is done it is going to enter into the second stage where the DNA coils around the DNA A multiple opiates which leads to the topological stress and once the topological stress is been generated it is actually during the presence of the in the presence of ATP the DNA A influences the 13 mer ATRH DUE region to start melting and once the DUE region is start melting the further melting is carried out by the recruiting hexamer protein which is called as helicase or the DNA B.

DNA B helicase clamps around each of the 2 single standard strands of the DUE site of the original C or C and the clamping of DNA U is supported by the clamp loader DNA C they make the DNA B DNA C complex and that is how the it is going to have the initial melting of the DNA at the original replication and that is how it is actually going to form the replication fork and the replication bubbles. Then we have the DNA C which is going to open up the DNA B ring and helps in placing the ring around the single standard DNA at the origin. While DNA B moves forward with the help of the ATP hydrolysis the single standard DNA binding protein covers the single standard DNA to prevent the unwinding and single standard DNA bind cooperatively in a sequence dependent manner sequence independent manner. So, single SSBs are actually going to bind the nucleotides and that is how they are actually going to destroy the affinity of the complementary strands and that is how it is actually going to keep the strands into the single standard DNA. Then we have the next is DNA B recruits the DNA G which is called as RNA primase to synthesize the RNA primer on both the both the strands which is called as leading strand or the lagging strands.

So, and then we are going to have the RNA primer which is influences the DNA C to release the DNA B from the site right and to initiate the elongation phase the primosome of prime formation occurs and as a result it is actually going to synthesize the primers and the primosome is actually a functional complex which is going to have the DNA G, DNA B and SSB and some of the accessory proteins. So, in the initiation exactly what will happen is that you are actually going to have the binding of the some of the components at the DU site and some of the component at the DUR site. Once these two sites are actually going to be occupied by the initiation factors DNA B and other kinds of proteins then you are going to have the recruitment of the helicase And helicase is actually going to be a hexagonal protein. So, it is actually going to run in both the direction. So, it is going to run in this direction and this will run in this direction and that is how it is actually going to form the replication bubble.

And once it is going to be formed it is going to allow the binding of the single standard

DNA binding protein and that is how it is actually going to allow the binding of the is how there will be a synthesis of the primers and one replication one fork will run in both the directions and that is how it is actually going to have the one under strand which is called going to be called as leading strand the other strand is going to be called as lagging strands. And once this initiation stage is over then it will enter into the elongation stage. So what is the basics of the elongations? So mainly the DNA pol III does not does the polymerization chain elongation happens by the free 3 prime OH primers attacking the alpha phosphoryl group of the incoming ATNTPs as a product byproduct of the reaction as mentioned earlier the pyrophosphate is going to be generated the bond formed between known is the phosphodiester bonds and that is how it is actually going to start adding the nucleotides based on the base pairing information. So once you have the Watson-Crick based pairing information which is going to be available through the DNA pol III it is going to be keep adding the nucleotides and these keep nucleotides are going to be coupled with each other by the phosphodiester linkage and the pyrophosphate is going to be released and this pyrophosphate is going to be get hydrolyzed by the pyrophosphatase and that is also going to generate the energy which is going to be utilized into the process. In this case there are two replication forks generated in for the prokaryotic that moves in the opposite direction for each other replication forks proceeds bidirectionally at a speed which is 1000 base pair per second per fork both the leading and the lagging strands are replicated simultaneously.

What is the leading strand? The DNA polymerase III synthesizes the strand from the 5 prime to 3 prime direction continuously towards the replication fork. So one fork is running in the direction of the 5 prime to 3 prime the other is running in the reverse direction that is the in the reverse order and that is what is called as lagging strand. So in the lagging strands the synthesis happen in the 5 prime to 3 prime direction but the disk continuously away from the replication fork. So if you have a bubble and if this is the bubble and if the replication is going in this direction machinery is going in this direction then this is going to be the leading strand and this is going to be lagging strand because it will wait for this region to be available to get open and then only it is actually going to do the synthesis in the reverse direction and that is why this is going to be called as the lagging strand and this is going to be called as the leading strands.

So there are three different stages of the elongations right. So DNAB helicase separate the two DNA molecules binding to the lagging strands template at the replication forks and moving along according to the polarity of 5 prime to 3 prime direction. Now DNAG the prime is associated with the DNAB and synthesizes the RNA primers complementary to the associated single standard strands. Interaction between the DNAB and DNAG regulates the Okazaki fragment length. Tighter association results in more frequent short fragments whereas the loose interaction will produce a longer fronted on lagging strands.

Length of the Okazaki fragments could vary between the 1000 to 2 herm nucleotides. Formation of replisome so it contains the DNA pol III holoenzyme and the associated protein like the DNAB and the DNAJ DNAG. This replisome starts the joining of the DNA piece by forming the phosphodiester bonds. DNA pol I removes the RNA primers of both the strands by the 5 prime to 3 prime exonuclease activity which generates a gap after the primer remover in the both the strands. DNA pol I also fill in the gap between the lagging strand fragment after the primer remover. Lastly the left over nicks are sealed by the help of DNA ligase with the help of the NAD⁺ plus as a energy source.

So this is what exactly going to happen in the replications elongation stage. So in the elongation stage what will happen is that you are going to have the replication fork which is moving in this direction. So this is going to be considered as the leading strand and this is going to be considered as lagging strand. So once this portion is going to open then the synthesis is going to start from this direction. So it is going to happen in this direction. Whereas in the case of leading strand because the synthesis always occurs in the direction of 5 prime to 3 prime and so it has to wait for the 3 prime to be available for making a primer and that is how it is actually going to be in the opposite direction.

And that is why this strand is going to be called as lagging strand and this strand is going to be called as leading strands. Now what is the role of gyrase in the elongation? So due to the unwinding of with the help of the DNA helicase the double stranded DNA in front of the fork become positively supercoiled. If the supercoiling increases the fork will

Therefore to overcome the halting from the supercoiling topoisomerase are needed. In E. coli the DNA gyrase the type 2 gyrase was discovered by the scientist called as Martin Gallet. And gyrase contains the two different subunits gyre A which contains cut and rejoin the DNA and gyre B which is responsible for providing the energy by the ATP hydrolysis. Then we also have the proofreading activity because when you are doing a DNA applications it is possible that the DNA polymerase could add the some nucleotide which may not be match which may not be as per the information available onto the template. And in that case there is a proofreading activity required right because when you are synthesizing a product it has to go into the into a testing stage or it is going to be get into a stage where you should test whether the product what you are synthesizing is of good quality or not. So proofreading activity is going to ensure that there is a sequence what is being produced is exactly identical of the template.

So DNA replication is amazingly accurate with one error in 1 billion nucleotides incorporated. This above mentioned tolerated mutant level is desired for the large

genome size especially. DNA polymerase carried out the process by their 3, 5 to 5 prime exonuclease activity. When an inaccurate nucleotide is incorporated the synthesis rates get reduced due to the wrong positioning of the 3 prime OH. This works like one delete key removing only the most recent error. So what happen is that when you are actually synthesizing and synthesizing it keep going and keep checking whether the attachment what I have made is actually making a pair with the template or not.

So if it is not then it is actually going to go back and it is going to do the corrections. Now the third process is called as the terminations. As there is a bidirectional replication the fork will melt at a position diametrically opposite to the ORC on the genome. Termination region contains multiple copies of 23 base pair long sequences or TIR sequences. Every TIR sequence act as a recognition site for a protein which is called as TUS.

So termination utilization substrate proteins. TUS-TIR complex allow one replication fork to pass if it is moving in a one direction but blocks the progression if it is from the opposite direction. The directionality depend on the TIRs protein localization on the DNA helix. In E. coli the orientation of the TUS-TIR complex is such that it ensure both the fork movement will be stopped at or near the same point.

After the complete replication process that the new two circles are physically interlocked or catenated. This decatenation is carried out by the topoisomerase IV to generate the separate two double standard DNA molecules. So termination is being done by the places where you are going to have the TIR sequences and these TIR sequences are going to be recognized by a protein which is called as the TUS and or the termination utilization substrate proteins and the TIRs and the TIR is going to form a complex allowing only one replication to fork to pass. But if there are two replication forks which are you know going to reach to that particular point and they will try to bypass then it is actually going to halt the replications. Now let us talk about the summary of the things what we have discussed so far right. So there will be and what will be the difference between the replication between the prokaryotic and as well as the eukaryotic system.

So in the initiation you are going to have the DNA A which is actually going to participate. So initially you are going to have the initiation reactions where you are going to have the DNA A whereas in the prokaryotes you are going to have the protein which is called as ORC. Then you also going to have the activity of helicase which is going to unwind the DNA and it is going to be DNA B then that is been done by the MCM complex in the eukaryotic system then you are going to have the helicase lower that is going to be DNA C whereas the same function is been done by the Cdc 6 and Cdt 1. You do not have to worry about all these proteins because when we are going to talk

about eukaryotic We will discuss all these. Then the single standard DNA binding proteins it is going to be prokaryotic SSBs whereas in the case of eukaryotes it is going to be RPA. Then we have the primase so it is going to be DNA G in the case of prokaryotic system whereas in the case of eukaryotic system you are going to have the pol alpha primase.

And then we have the polymerase so you are going to have the DNA pol III which is the main polymerase required for the replication in the prokaryotic system whereas in the case of eukaryotes you are going to have the DNA pol III delta and the poly DNA pol epsilon. Clamp you are going to have the beta clamp whereas in the case of the prokaryotes you are going to have the PCNA ring. Then we have the clamp loader so you are going to have the gamma complex whereas in the case of the eukaryotes you are going to have the RFC. Then we have the ligase, ligase is required for you know joining the lagging strands right.

So it is going to be a DNA ligase whereas in the case of eukaryotes going to be a DNA ligase I. Then we require the primase removal so primase removal is being done by the DNA pol I or ribonucleus H in the case of prokaryotes whereas in the case of the eukaryotes it is going to be done by the RNase H or FEN I. So these are the some of the components what are being different between the prokaryotic system and the eukaryotic system. And what we have discussed so far we have discussed about the DNA replication machinery in the prokaryotic system. We have discussed about the origin of the applications, features of the origin of the application and so on. So in a subsequent lecture we are going to discuss some more aspects related to the prokaryotic application. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 05
Replication

Lecture-22 Replication in Prokaryotic System (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati and what we were discussing? We were discussing about the different aspects of the molecular biology in the course in this particular course. So, in this particular module we are discussing about the central dogma of molecular biology and in this context so far what we have discussed? We have discussed about the replications and so we have divided the discussion about the replication that it is going to be first we are going to discuss about the replication in prokaryotes and then followed by that we are going to discuss about the replication in eukaryotes because replication the sole purpose of replication is that it should synthesize the genomic content of that particular organisms in a limited time period so that it should be able to divide and should be able to multiply its number. So, what we have discussed so far? We have discussed about the replication in the prokaryotic system and what we have discussed that it is going to have the origin of replications and original replication is going to be a AT reach sequences so that it is easy for the helicases and other enzymes what is been involved into the active into the replications and they should be able to melt the DNA and should be able to unwind the DNA very easily. And then followed by that we have also discussed in extensively about the machinery what is required for the DNA replications in the prokaryotes. We have discussed about the DNA pol 1, pol 2 and pol 3.

We have discussed about their structures functions and all other kinds of features and then apart from that we have also discussed about the helicases, DNA B, SSP structures, topoisomerases and other kinds of enzymes. So, in the previous lecture we have discussed all these aspects. Now in the current lecture we are going to discuss about the special type of replications and special type of replications what is been found into the prokaryotic system and how what are the different what are the differences and what are the different machinery what is required for performing these type of this type of replications in the organisms. So, we have discussed about the replications and we have discussed about the replication in prokaryotes.

Now we are going to see the special mode of replications what are been available or what are been found into the prokaryotic system and how what kind of adaptation is happening into the bacteria and how the different type of cellular machinery is involved. So, replication if you see the replications replication has the very distinct three steps

right. You are going to have the initiation followed by the elongation and followed by the termination. So, if you recall we have discussed that the within the initiation the cellular machinery will identify first identify the origin of replication and there would be only one single origin of replication into the prokaryotic system whereas, there are going to be multiple origin of replication into the eukaryotic system. And once they have identified the origin of replication then the initiation complexes are going to be formed and the helicases and other kinds of enzymes will go and sit on to the initiation site and that is how they are actually going to unwind the DNA.

While they are unwinding the DNA they will also going to ensure that the DNA what they are unwinding and preparing it for the replication is methylated if the DNA is hemimethylated or unmethylated then the they will not actually going to initiate the replications. After that it will enter into the elongation phase where the DNA polymerase will go and sit and then start adding the nucleotides as per the borson-cirque base pairing structures and so on. And then it will enter into the terminations termination and the termination is going to termination is being done by the terse sequences which are going to interact with the termination machinery and that is how it is going to terminate. So, these are the some of the basic structures or basic mechanism what is happening, but in some of the bacterial system you are actually going to have the replication in a special mode. So, one of the such special mode is called as the rolling circle replication or rolling circle model.

So, in a rolling circle model it is been found in the archaeabacteria, bacteriophage, plasmids and the viral DNA such as the HHV, HPV, Gemini virus and the viral RNAs. In a rolling circle model what you have is you have a double standard circular DNA and this is actually going to have the so there will be a nick what is going to be formed onto the DNA and then the machinery will actually going to utilize the information from the inner strands and that is how there will be a synthesis of the DNA on one end. And it will be utilizing this information keep rolling like this and that is how they are actually going to start synthesizing. Once they will synthesize the one copy of genome then it is going to be cut and then this is actually going to serve as a template to synthesize the second strands. So, what are the different enzymes what are going to play a crucial role into the rolling circle model.

So, you are going to have the RepA. So, RepA is going to initiate and nick the double standard origins or DSO. Remember that this is the first event that you are going to have you have to first form a nick structure. So, that the machinery will enter into this structure and it is going to utilize the nucleotide what is present onto the inner circle and that is how it is they are going to synthesize the this strand and that is how it is actually going to make the multiple copies. Then you require the PCRA or plasmid copy reduced

and this is going to be a helicase which moves the nicks strands.

Then you require the DNA pol III and DNA pol III is going to have the 5 prime 3 prime polymerase activity and it is going to be male replicates. Then you require the DNA pol I and DNA pol I the function of the DNA pol I is going to be that it is going to remove the RNA primers and then you also require the DNA ligase. DNA ligase is going to join the end to make the strand circular. Now these are the some of the steps what you are going to perform. So, in the initiation, elongation and termination.

So in the initiation the RepA is actually going to recognize the double standard origin. So, in this case you are going to have the for example, you are going to have the circle and then in this suppose this is the place. So, this is going to be a place where you are going to have the double standard origin or DSO. So, this is actually going to go and allow the sitting of the RepA and after recognition it is actually going to make a nick. So, it is going to make a nick into the upper strands.

So, what is mean by the nick is it is actually going to make a break actually. Then in the elongation phase the RepA stay attached with the nicked 5 prime phosphate end. So, it is actually going to make a cleavage and it is going to hold the DNA. Whereas the 3 prime end is 3 acting as the DNA pol III primers and the DNA pol III act in a 5 prime to 3 prime direction produces the multiple concatamer multiple single standard copies of the original DNA series to increase the efficiency of all 3, 1 replicates PCRA is incorporated before it unwind to the double standard DNA. And then in the termination state you are going to have the RepA which is attached cut the leading strand to stop the replication of that strand.

Therefore, the second strand is to be left to synthesize to replicate the second strand the RNA polymerase or the primer synthesis RNA polymerase and DNA pol III copies the single standard origin and elongates after replication the DNA pol I remove the RNA primer and put the correct bases there and then the DNA ligase will come and it will join the nicks to give you the double standard circular DNA molecule. So, this is exactly what is going to happen. You have this is the single standard original replication and this is going to be the double standard original replication. So, when you have the so in the initiation stage your RepA is actually go and come and bind to this particular portion and then it is actually going to make a nick. After making a nick the RepA is actually going to bind the 5 prime end whereas the 3 prime end is going to be free.

And then on the 3 prime end the pol III is actually going to bind and then it is actually going to utilize this portion as a primer and that is how it is actually going to start synthesizing the DNA and that is how it is actually going to make the same copy, the

single standard same copy right. So, it is actually going to make the single standard copy so it is going to run like this And that is how it is actually going to make the single standard copy. So one strand is going to be synthesized fully. So for example, it goes like this and one single copy is going to be synthesized and then Rep A is going to cut the strand and DNA ligase is going to join the links. So it is going to have the double standard and it is going to have the other double standard DNA molecules and the single standard DNA molecule is produced which will be synthesized in a double standard molecule using the similar kind of steps.

So it is very clear that in the double standard origin of applications the Rep A is going to bind and then it is going to make the nick and that nick is going to serve as a primer for the DNA pol III and then the DNA pol III is going to synthesize and that is how it is actually going to make the single standard concatamer. These concatamer are going to be cut right once they will one round is over right and that is going to be double standard simply by the DNA pol I and that is how it is actually going to have the double standard circular DNA molecule the daughter DNA molecule and the other molecule is also going to be utilized. So in this one you are going to have the two strands right this strand and this strand. This strand is going to serve the template to synthesize the outer strands and afterwards once the application is over then the outer strand is also going to be utilized to synthesize the inner strands and that is how it is actually going to complete the cycle. One of the classical example of the ruling circle model is that it is going to be present in the some of the E.

coli species right. So you are going to have the donor cell you are going to have the recipient cells and once they will going to form there will be a conjugation and then after the conjugation they are going to exchange the genetic material. So how they are going to exchange genetic material is that the donor cell is actually going to have the ruling circuit model right and because of that its genome is going to be replicated and it is going to have the concavitabra right. So this single standard DNA is going to be transferred to the next recipient DNA and that is going to be present to the recipient cells and then the circular and the double standard plasmid is produced by the new cells. So this single standard circular DNA is going to be then utilized as a template by the recipient cells and it is going to make the double standard DNA and that is how you are actually going to have the plasmid and that is how you are actually going to have the exchange of genetic material and that is how you are going to exchange the phenotypic feature. So for example if this bacteria is ampicillin resistance right and this bacteria is ampicillin sensitive then this bacteria is actually going to provide that resistance through the ruling circuit model and provide the DNA responsible for that and this DNA is going to provide the ampicillin resistance even into the donor molecule also.

So apart from the ruling circuit model you are going to have the another kind of replication mode that is called as D loop formation or the D loop replications. This replication is found in the small circular and organellar DNA for example it is present in the chloroplast and mitochondria. Remember that the chloroplast and mitochondria although they are present within the eukaryotic cell but they are not eukaryotic in origin right they are already they are being prokaryotic in origin. So they follow many of these features what is present in the bacterial cell right for example they have the circular DNA they have their own DNA. So chloroplast and mitochondria are actually going to follow the mechanism what is being known for the or whatever we have discussed so far into the prokaryotic replication system right.

So where one triple standard structure called a displacement loop is going to be formed and mitochondrial DNA it is actually the 16.6 kB which is consist of two strands heavy strand and the light strand. So this is the heavy strand the inner strand and the outside is the light strands. It comprises a lengthier NCR or the non-coding region and act as the regulatory region. In this region the mitochondrial DNA has its promoter for the transcription.

One is the light strand promoter and the other is called as the heavy strand promoters or HSP. Conserved sequence MOXES CSBs and the termination associated sequences are also present. How the replication terminate at the task is still not known. So NCR also contains the origin of leading or the heavy strands. So this heavy strand is called as the leading strand whereas the light strand is also called as the lagging strand.

So origin of leading strand OH is on this heavy strand whereas the origin of lagging strand is written onto the light strands. Now what are the different key players which are involved into the D-loop replications. So you are going to have the D-N-A-pol gamma. So it is a main replicates or the polymerase. It has two subunit polymerase A or PolA.

It is from the DNA pol of the BA family. It shows the proofreading activity in the 3 to 5 exonuclease activity and it is highly accurate. That is the one error in the 1 billion base pair. Then we have the PolB which has the add-on subunit to improve the interaction between the PolA and the DNA templates. It increases both catalytic activity and the processivity.

Then you also require the Twinkle. So Twinkle is a hexamer helicase which needs the fork structure for being loaded and for start unwinding. Then you require the mitochondrial single standard DNA binding protein. So SSBs binds with the newly formed single standard DNA to protect from the nucleases and maintain the single standard structures.

Then you also require the PolRMT. So mitochondrial RNA polymerase synthesizes the RNA primer onto the displayed strand as it cannot work on the single standard DNA as a template. Then you also require the DNA ligase 3. So DNA ligase joins the nicks in the new DNA strands. Then you also require the RNase H1 and MgM1. So mitochondrial genome maintenance exonuclease 1 that helps into the primer removal after the replications.

So these are the steps of the D-loop formations, D-loop replications. So you start at the OH site. The PolRMT starts making the primer onto the H strand to synthesize the whole H strand. So you are going to have the genomic content in the form of the two circular strands. You are going to have H strand which is on the blue one and you are going to have L strand which is the red one.

So in the first part on the OH side which is present onto the H strand is PolRMT is going to sit and it is going to start replicating that. Whole parental displays H strand is covered by the single standard DNA binding protein. The binding of the SSB stop the PolRMT for a random RNA synthesis onto the H strands. So this is what is going to happen here. Then the twinkle the helicase comes before the Pol gamma to unwind the double standard DNA to move the fork while the SSB maintain the single standard DNA leading to the formation of the replisome.

So the replisome is going to be formed onto the DNA strand. After the two third of the genome is synthesized the replisome passes through the OL side or the origin of the application onto the light strand or L strand actually. There the parental single standard H strand forms a stem loop kind of structures. So this is a stem loop of structure which is going to be formed. The stem loop structure stops SSB from binding thereby the one short stretch of the single standard DNA in the loop become accessible for PolRMT to make the RNA primers.

So once you have this it is actually going to stop the binding of the SSB and the other hand it is actually going to allow the PolRMT or it is a RNA polymerase to synthesize the primer for the L strand also. And then after 25 nucleotides of stem loop structure the polymerase gamma replaces the PolRMT at a 3 prime end of the primer which results into the synthesis of the L strands. H and L strands are replicated continuously until they reach to the termination sequences and forming a triple strand displacement loop D loop and then the two DNA strands are going to be formed. So this is what it is going to happen here that you are going to have the L strand and H strand and in the initial region the L strand on the L strands of the origin of the application that is the O H side the PolRMT is going to bind and that is how it is actually going to start the applications.

Whereas on the L side first it is actually going to have the stem loop structure and this stem loop structure is going to stop the binding of the SSBs and it is also going to serve as a primer and that is how it is actually going to allow the binding of the PolRMT and it is going to synthesize the primer and that is how it is actually going to initiate the synthesis of the L strands and ultimately the both the L strand dimer or the H strand is actually going to be formed and that is how you are going to have the two double standard daughter DNA.

So from the this parent DNA you are going to have the two daughter DNA at this end. So this is all about the DNA replications into the prokaryotic system what we have discussed so far we have discussed about the origin of the applications into the prokaryotic system and this origin of the application is has a classical features of ATDG regions and it is going to be recognized by the cellular machinery and that is how the helicases and other cellular proteins are actually going to go and bind and unwind the DNA and then it is going to be enter into the elongation phase and in the elongation phase the DNA polymerase is going to synthesize the DNA strands and it is going to join the nucleotide onto the incoming 3 prime site and that is how it is actually going to synthesize and once it reach to the termination sites it is going to be terminated by the TIR sequences and so on. And in the current lecture we have also discussed about the special mode of DNA replications so where we have discussed about the rolling circle model which is very very common in the bacterial system and then we also discuss about the D loop formation which is more common in the in the organelle DNA such as the mitochondrial DNA and the chloroplast DNA. So with this small discussion about the prokaryotic replications I would like to conclude the lecture here in our subsequent lecture we are going to discuss about the DNA replications in the eukaryotic system. Thank you. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 05
Replication
Lecture-23 Replication in Eukaryotic System

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati and this particular module we are discussing about the DNA applications. So, far what we have discussed we have discussed about the DNA replication in prokaryotes and while we were discussing about the DNA replication in prokaryotes we discussed about how the initiation complex is going to be formed and how the different types of components are required for the DNA replication in prokaryotes and how the DNA replication in prokaryote is different and required the special machinery for its DNA synthesis. Now in today's lecture we are going to discuss the about the eukaryotic replications because eukaryotic replication is different from the prokaryotic replication in terms of the components or the enzymes what is required and the property of these enzymes are very different. So, as a name suggests the DNA replication is the mechanism by which the cell duplicates its genetic material ensuring that the each newly formed cell receive an accurate copy of the original DNA. So, this kind of objective is also being met even when we are doing the DNA replication in the eukaryote as well.

The process is essential for the growth, development and transmission of genetic information from the one generation to the next generation and the DNA replication involves the various enzymes and protein working together to unwind and separate the double standard DNA molecule synthesize the new complementary strand and ensure the fidelity of the copied genetic material. The basic principle of eukaryotic and prokaryotic replication are the same however there are some notable differences. Now the one important point is that the eukaryotic replication is much more complicated than the eukaryotic replications and there are many reason why it is so actually. So, why the eukaryotic replication is more complex than the prokaryotic replication because of the simple reason that the eukaryotic genomes are quite complex they are larger than the bacterial DNA.

Remember that the bacterial DNA's are very small even in the plasmid DNA also and it is having a complex structure because remember that the bacterial chromosome or bacterial genomic content is not been associated with the protein molecules and it is not that much complex compared to that the eukaryotic system. Remember that when we were discussing about the eukaryotic genome we discussed that how it is actually been

associated with the histone proteins and how the histone octamers are forming the nucleosomes and then nucleosomes are further assembled and give you the chromosomes. So, the chromosome is mostly present in the eukaryotic structures. But apart from these differences or the complexity the replication process in the both the prokaryotic and the eukaryotic system actually involves some of the basic steps. For example, the formation of the replication fork is common between them, primer synthesis is also common between them.

Once the primer job is over and it has to be removed by the DNA pol I in the prokaryotic system that is also common then you also going to have the okazaki fragment whether it is the prokaryotic system or the eukaryotic system then the replication mode it is going to be semi conservative whether it is a eukaryotic or prokaryotic system and the movement of the replication fork would be bi-directional in the both the prokaryotic as well as the eukaryotic system. So, replication fork is going to be bi-directional in the case of prokaryotic or eukaryotic system and it is also been required for gap bridging between the newly synthesized DNA fragments with the help of the DNA pol I and the DNA ligase. So, apart from these kind of similarity there are significant difference in terms of the machinery and other kinds of proper requirements. So, apart from these similarities there are significant difference between the prokaryotic and the eukaryotic replications. We have discussed many of these differences so but I thought I should remind you so that it will be easy for you to follow up the eukaryotic replications.

So, what is the difference? The important differences between the prokaryotic and the eukaryotic replications. Eukaryotic DNA is larger than the prokaryotic DNA that is a very very important difference because the purpose of the replication is to duplicate the DNA and if the DNA amount is large then it will actually going to be required the different machinery. The movement of fork is slower in eukaryote than the prokaryote because the nucleotide have to disassemble so that the DNA becomes available to the DNA polymerase. So, this means the eukaryotic replication is going to be slower it is going to take longer period of time to complete and that is why you remember that the life cycle of the eukaryotic cells are much more you know much more than the prokaryotic system. For example, the E.

coli completes its life cycle or completes its duplication in the 18 minutes. Compared to that a simple e coli simple mammalian cells for example, the He k 293 completes its duplication somewhere around 16 to 17 hours. So, it that is because of the simple reason that the DNA replication is very very slow in the case of mammalian system compared to the bacterial system. So, here everything is getting over by in 18 minutes whereas, here that same thing required the 16 to 17 hours. Then it has a distinct packaging of the

eukaryotic DNA in terms of the chromatin.

So, chromatin structure is very very higher order organizations in the case of the eukaryotic system whereas, it is not in the case of eukaryotic system. Now, let us take an example like what is the replication rate. So, replication rate in the case of prokaryote is approximately 1000 base pair per second whereas, the replication rate is 10 times slower than the prokaryote. So, it is only the 50 nucleotides per second. So, it is very very small and it is a small because of the simple reason that the nucleotide have to disassemble before the DNA become available to the DNA polymerase.

And apart from that in the prokaryotic system you do not have the chromatin structure. So, chromatin structure has to you know disassemble the DNA has to be free from the chromatin structures and so on. So, before getting into the detail of the different processes of the DNA replication in the eukaryotes let us first understand the machinery about the DNA replications. So, there are different types of polymerases because these are the important component of the replications. So, you have the 5 different types of DNA polymerases like alpha, beta, gamma, delta and the epsilon.

And the localization most of the DNA polymerases are present in the nucleus except that the DNA poly gamma is present into the mitochondria. Then the what is the biological function? The biological function of the alpha is the replication initiation, beta is involved in the DNA repair, then the gamma is involved into the mitochondrial DNA replications, then the delta and the epsilon is required for the replication of the lagging as well as the leading strands. Then you require the number of subunits. So, as far as the structure is concerned the alpha is tetramer, beta is monomer and gamma is the homotetramer and the delta is dimer and the epsilon is not known. Then 3 prime exonuclease activity, 3 prime exonuclease activity is absent in most in the alpha and beta, but it is present in the gamma, delta and epsilon.

Then primase binding, so primase binding is present in the case of alpha, but it is absent in the all other DNA polymerases. Then you require the molecular weight of catalytic site. So, it is going to be 160 to 185 and so on. Then you require the Km for the dNTPs. So, Km for the dNTPs is in the range of 2 to 5, 10 and 0.

5 and 2 to 4. So, if you cannot be able to understand what is Km, you should you know be able to understand this by I think we have discussed very briefly in one of the lectures where we were talking about the enzyme. So, we have discussed about the Km and so Km is the Michaelis Ventum constant and it actually indirectly says what will be the affinity of the enzyme for the dNTPs. Sensitivity to arabinol CTP, so it is going to be very high in the case of the alpha and delta, but it is going to be low in the case of beta

and gamma and sensitivity to Fe2 code line. So it is going to be high in the case of alpha and the delta where and the epsilon as well, but it is going to be low in the case of the beta and gamma. Now, let us talk about the some of these DNA polymerases.

So, we will start with the alpha. So DNA polymerase alpha is localized into the nucleus and it is a tetramer. So it is going to have the 4 subunits, you are going to have the PolA1, PolA1 regulatory, PolA3 primase activity and PolA4 it is a primase activity. So for all the 4 subunits are different. So A1 is going to have the catalytic activity, A2 is going to have the regulatory activity and A3 is going to have the primase activity and A4 is also going to have the primase activity.

Now that means you are going to have the Pol alpha and Pol primase activity together. This means it is actually going to synthesize the primer and as well as it is going to sit on utilize that primer for the DNA applications. So initiation of the replication on both the leading and lagging strand synthesis is the function of the DNA polymerase alpha. Then we have the DNA polymerase delta. So DNA polymerase delta is localized into the nucleus, it catalyzes the synthesis of the lagging strands.

It has the high processivity when interacting with the PCNA or PCNA or the polyfluidating cell nuclear antigen. PCNA is a important factor what is been having a very critical role in the DNA damage and repair as well and it is also associated with the helicase activity and it improves fidelity or replication by a factor of 10 to power 2 due to its proofreading actions. It has a 4 subunit, it has 1, 2 and 3 and 4. So large subunit catalyzes the 5 prime to 3 prime catalytic activity whereas small subunit catalyzes the 3 to 5 prime exonuclease activity or the proofreading activity. Then we have the DNA polymerase epsilon.

So it is localized into the nucleus. DNA polymerase epsilon catalyzes the repair mechanism. It also catalyzes the removal of primer and filling the primer gap in the Okazaki fragment. So it is very much close to what the function what you have seen in the case of prokaryotic system as DNA pol I. So it is actually going to have the same kind of role that it is going to remove the primers and it is also going to fill the gap in the between the Okazaki fragments.

It is going to have the 4 subunit. You are going to have 1, 2, 3 and 4 and 5 prime to 3 prime polymerase activity, 5 prime to 3 prime exonuclease activity and 3 prime to 5 prime exonuclease activity is present in the DNA polymerase epsilon and it is required for the different types of activities. For example 3 to 5 prime to 3 prime exonuclease activity is required to remove the RNA primer whereas 3 to 5 to 5 prime exonuclease activity is required for the proofreading. Now what is the replication factor RFA or

replication factor protein RFA? So it plays a significant role in stabilizing the single standard DNA region that are exposed during the DNA replication and repair mechanism. RPA prevents this single standard region from forming the secondary structure and protect them from the degradation allowing the other enzyme and factor to perform their function accurately.

So RPA is actually going to do the same job what you have understood in the case of SSP actually into the prokaryotic system. And then we have the PCNA. PCNA is important for the DNA synthesis and the repair and we are going to do discuss in detail about its role into the DNA repair when we are just going to discuss about the DNA repair mechanisms. The molecular weight of the molecule is 35000 kilo Dalton it is a multimeric protein and it is found into the large amount to the nuclei of the polyfluidating cells. And what is the function? So PCNA act as a clamp to keep the DNA polytheta and delta from dissociating off from the leading strand and PCNA help both hold the DNA polymerase epsilon to the DNA.

Replication factor C or RFC also known as clamp holder or to the matchmaker. So this is the PCNA which is going to be a clamp and then it is going to be a RFC so it is going to make a complex with each other. So binding of PCNA and RFC is going to make a complex and this complex is going to have the affinity for the DNA. So binding and hydrolysis of ATP once this is formed it is actually going to bind the ATP and it is going to hydrolyze the ATP and that actually is going to bring the structural changes into the PCNA and clamp holder RFC. And once it there will be a structural changes into the RFC it is actually going to have the affinity for the DNA and that is how it will go and bind to the DNA.

And then that actually is going to load the PCNA onto the DNA and once it binds to the DNA then it is actually going to form a complex with the polytheta and there will be hydrolysis of the ATP and the DNA polymerase delta is going to be recruited onto the DNA poly on the DNA and that is how it is actually going to help in the initiation stage of the DNA applications. Now one of the important aspect of the DNA replication in the eukaryotes is that it is actually the DNA is not freely available compared to the prokaryotes where the DNA is freely available and it is only required to locate the origin of the application and then all the machinery is going to assemble onto the original replications and then it is actually going to start the synthesis. Compared to that here first you have to bring the free double standard structure and then you are actually going to unwind the DNA and then you are actually going to do all that what you have discussed in the prokaryotic system. So, the first thing is you have to dismantle the chromatin. So, that a chromatin should be available for further that so, that the free DNA is available for all these kind of activity.

So, histone dissociation and then associations and all these events has to be reversed once your replication is done otherwise this free DNA which is not covered with the protein would be vulnerable for the different types of DNases and other kinds of enzyme what is present inside the nucleus. So, DNA replication is sandwiched between the two additional steps in the eukaryotes dissociation of the histones on the synthesis of histone. So, methylation at the fifth position of cysteine residue by the DNA methyl transferase appear to be functioned by loosening up the chromatin structures. This allows the DNA access to the protein and enzyme needed for the DNA applications. Remember that the DNA and the histones are attached with each other by a positive negative interaction.

So, it because of this electrostatic interactions. So, once the histones once the cytosine in the DNA is going to be methylated by the transferases it is actually going to bring or it is going to loosen the interaction between the histone and the chromatin structures and that is how it is going to allow the DNA access to the protein and enzyme needed for the DNA applications. Then it occurs simultaneously with the DNA. So, synthesis of the histone it is occurs simultaneously with the DNA applications. So, these are the sequential steps into the DNA applications you are going to have the first step is the formation of the reinitiation complex.

The second step is the initiation the third step is the elongation the fourth step is the termination and the fifth step is the telomerase function. So, that you can actually be able to have the completion of the telomeric regions. So, the first start with the pre initiation the pre initiation step is the crucial step that prepare the DNA for the actual replication process. This is steps primarily occur at the origin of replication which are specific DNA sequence where the replication begins. The process of identifying these sequences is known as the replication replicator selections which occurs into the G1 phase.

This process leads to the assembly of multi protein complexes at each replicator in the genome and the origin activator only occurs after cell enter into the S phase and trigger the replicator associated protein complex to initiate the DNA unwinding and the DNA polymerase recruitment. So pre initiation complex is come is a is a is been formed on to the origin of the applications. So the combination of the ORC MCM2 to 7 and CD66 and CDT1 along with the other regulatory protein factors form the pre initiation complex at the origin. This complex serves as a platform for the initiation of the application. So, imagine that if this is the origin or applications then the ORG will go and bind and once the ORG will go ORC will go and bind then the CDC6 and CDT11 is actually going to bind to this particular ORC.

And once these are going to bind then you are going to have the binding of the MCM2 to 7 and these are the once these are going to bind it is actually going to make the pre initiation complex and these pre initiation complex job pre initiation complex is that it should allow the recruitment of the DNA polymerase so that it will actually going to start the DNA synthesis. So, these are the some of the crucial step into the pre initiation complex formation. Now you are going to have the initiation. So, it involves the coordinated action of various protein complexes and enzyme to ensure the accurate and faithful duplication of genetic material. This process ensure that the each daughter cell receives a complete copy of the genome during cell divisions.

So, you are going to have the autonomous replicating sequences or the ARS or the replicators. For example, each contains approximately 400 automatic replicating sequences. So, these automatic replicated sequences are the independent sequences they are actually going to have their own origin or application. So, they can be able to you know start the replications and remember that in the case compared to the prokaryotic system in the eukaryotic system you are going to have the multiple origin or application that is how you are actually going to be complete the duplication of the genomic DNA at multiple points. So, you are going to start like the DNA replication will start from one end of the DNA and then it is going to over start from there and then it is going to finish by the end of the DNA.

No, it is not like that. In the case of eukaryotic system the DNA replication is going to be start at multiple points and that you know because the replication rate is very low compared to the prokaryotic system. So, it is actually required the multiple points at which the DNA replication is going to start. The second point is because the DNA size is very big it needs the multiple machinery to replicate the things. So, a specific site for the initiation of the DNA replication is the AT-rich sequences which is highly conserved 11 base pair sequences then you also have the flanking sequences and then you also require the 100 to 150 base pair long the 3 prime ends sequences. So, it may be the AT-rich sequences are actually the original replications the sequences or the site where the pre initiation complex is going to assemble and then the initiation is going to start.

The multiple origin of replications are spaced from the 300 to 300 kB apart which means for example if you have a DNA then you are going to have the multiple origin of the applications. So, all these original applications would be somewhere around 300 kB which means from this particular original application one fork will run in this direction and another fork will run in this direction and that is how it is actually going to complete the replication or the duplication of this amount of DNA. So, for example if it happens up to this so this is at this from this origin of the application it is only going to give you a DNA until this. But then you are going to have another original replication that also is

going to run in this direction and this direction and that also going to synthesize this amount of DNA. So, in a same amount of time this original replication will give you the DNA number 1 this will go actually going to give you DNA number 2 and this again this will going to give you the DNA number 3 and so on and that is how these all are actually going to assembled later on and it will give you the complete synthesis of this particular whole stretch.

So, that is what the adaptation or that is what the stretch D is what going to be adopted by the eukaryotic system because the eukaryotic genomes are very very large compared to the bacterial genome. The sequence between the 2 original replication is known as the replicons. So, this is actually a replicons this is the one replicons which is actually going to be you know participate into the replications this is another replicons this is another replicons. So, these are the multiple replicons what are going to be formed into the eukaryotic system. The AT rich also known as the ARS or the automatic replicating sequences similar to is similar to AT rich 13 mer present in the E coli OEC.

It is also called as the ORE or the origin replicating elements. The flanking sequences consist of the overlapping sequences that include the variant of the core sequences. So, ORE or the ORC so, ORE is called as origin replicating elements and ORC is called as origin replication complexes. So, ORE which is a 111 base pair sequence in the core sequence bind to a set of proteins for example, DNA pol alpha helicases DNA pol delta RFC, PCNA SSB, RFA and that all are going to assemble on to the origin recognition complexes and all these are going to make the origin recognition complexes which is a multimeric proteins. And initiation of the replication in all eukaryotic require this multimeric protein which binds to the several sequences.

So, ORE located adjacent to the approximately AT base pair AT rich sequence that is very easy to unwind the binding of ORC to ORE causes the unwinding at the DU that is the DNA unwinding elements. Now, events in the replication fork the DNA synthesis is initiated by the ORC and ORE the replication forks move bidirectionally and replication proceeds simultaneously as many as 200 forks which means you are going to have the 200 origin of replications or replicons simultaneously for working at together. Then the formation of the replication forks the replication fork in the eukaryotic consists of 4 components that forms in the following sequences. Sequence number 1 the DNA helicase and DNA pol alpha unwind the short segment of the parental DNA at AT base pair AT rich sequences called DU or the DNA unwinding elements. Then the DNA pol alpha initiate the synthesis of the RNA primer which is going to be a 110 base pair RNA primers.

Then the daughter strand synthesis is initiated by the DNA pol epsilon and the DNA pol

delta in the leading strand respectively. So this DNA pol epsilon and DNA pol Hidah is going to have the initiation of the DNA in the leading strands SSB and RFA binds to the single standard DNA and prevents its reannealing. So DNA pol epsilon and DNA pol delta is going to have the initiation strand is DNA synthesis going to initiate into the leading as well as the lagging strands. Then the 2 additional factors which play important role into the replication of eukaryotes are the PCNA and the RFC. So PCNA is actually going to be proliferating cell and nuclear integers and it act as a clamp to keep the DNA pol delta to keep dissociating of the leading strand and thus increasing the processivity of DNA pol epsilon.

Whereas the RFC is going to work as a clamp loader or matchmaker and its function is that it assists the DNA pol delta to form the clamp between the DNA and the PCNA and it helps in setting up a link between the DNA pol delta and DNA pol epsilon so that the leading and lagging strands synthesis can take place simultaneously. So this is the one of the examples where the fork is running in the both directions and this is the situation in the how the replication fork is going to be formed and so the replication initiation complex is going to be assembled on both the side what you see here is one side one fork is one initiation complex is going to be assembled on this side and another one is going to be assembled on this side and that is how it is actually keep you know removing the association of the DNA from the nucleosome and that is how the this will run in this direction and this will run in this direction. Rate of the replication fork movement so the rate of replication fork movement in eukaryote is approximately 50 nucleotide per second which is only one tenth of the eukaryote replication rate replication of human chromosome proceeds bi-directionally from the multiple origin spaced 300 to 30 to 300 kBBS pair apart and completed within an hour and average chromosome contain nearly 100 replicons between and thus the replication proceeds simultaneously as many as 200 replicons. So this is all about the pre initiation complex and the initiation once the initiation is done then it is actually going to enter into the next phase and that is called as the elongation. So elongation during elongation an enzyme called DNA polymerase add the DNA nucleotide to the C prime end of the newly synthesized poly nucleotide strands the template strand specify which of the four nucleotide that is ATGC is going to be added at the position along the new chain.

So you know that the wherever you in the template is the template has A then it is actually going to add the T if the template has G then it is actually going to add the C. So it is always going to follow the Watson-Crick base pairing rule and that is how it is actually going to add. Only the nucleotide complementary to the template nucleotide at the position is added to the new strands. For example when the DNA polymerase meet an adenine nucleotide onto the template strand it adds the thymine to the C prime end of the newly synthesized strand and then move to the next nucleotide on the template

strand. The above process will continue until the DNA polymerase reaches add to the end of the template strands.

So these are the some of the events what is going to happen in the elongation into the eukaryotic DNA applications. You are going to have the assembly of the DNA pol delta and the epsilon onto the leading and the lagging strands and that is how you are going to have the synthesis of the leading and the lagging strands. So you are going to have the recruitment of the polymerase and primases onto the both strands right. So this is going to be leading strand this is going to be lagging strand and then this clamp is actually going to keep sliding into this direction and that is how it is actually going to be keep unwinding the DNA and same is true for this one also right. And in this one you are going to have the synthesis of the lagging strands and this side you are going to have the synthesis of the lagging strand whereas on this side you are going to have the synthesis of leading strands.

Now for one of the important component of this whole reaction is the synthesis of the primer and that is being done by the enzyme which is called as primase. So primase all newly synthesized nucleotides strand must be initiated by the specialized RNA polymerase called as a primase. It initiate the poly nucleotide synthesis by creating a short RNA nucleotide strand complementary to the template DNA strands. The slot set of RNA nucleotide is known as the primers. Once the RNA has been extended at the template strand the primer exist and the DNA polymerase strand the new strand with the nucleotide complementary to the template strand.

RNA nucleotide in the primers are removed by the deononucleotide by the help of the DNA polymerase. Once the DNA replication is finished the daughter molecules are made entirely of continuous DNA strand with no RNA portions. The leading and the lagging strands the DNA polymerase can only synthesize new strand in the 5 prime to 3 prime direction so that the two newly synthesized strand grow in the opposite direction because of the template strand at the each replication force are anti-prolaryl. Leading strand it is synthesized continuously towards the replication fork as helicase unwind the template on double standard DNA. Whereas for the lagging strand it is synthesized in the direction away from the replication fork and away from the DNA helicase unwind.

Its synthesis this strand is synthesized in pieces known as the Okazaki fragment and each fragment began its own RNA primer. This all we have discussed when we were discussing about the prokaryotic system. Then we have the leading strand synthesis. So in the leading strand synthesis you are going to have it is initiated upon the RNA primer synthesis by the primase unit of the DNA pol alpha. Then the DNA pol alpha adds a stretch of DNA to the primers at this point the RFC carried out a process known as the

polymerase switching and RFC removed DNA pol alpha and assemble the PCNA in the region of primer strand terminus.

Then the DNA pol epsilon bind to the PSNA and carried out leading strand synthesis due to the 5 prime 3 prime polymerase activity. After the addition of several nucleotide in the total strand is removed by the DNA pol epsilon due to its 3 to 5 micro nucleus activity and gap is also filled by the same polymerase again. Then the nick is sealed by the DNA ligase and finally, the fidelity of the replication is removed by the DNA pol delta due to its proofreading activity. Then we have the lagging strand synthesis. The lagging strand synthesis of the Okazaki fragments initiated same way as the leading strand synthesis.

RNA primer is synthesized by DNA pol alpha due to its primase activity. The primer is then extended by the DNA pol delta due to its 5 prime to 3 prime polymerase activity using the dNTPs. All but one of the ribonucleotide in RNA primer is removed by the RNase H1. Then the exonuclease activity of FEN and the RTH1 complex removes the one remaining nucleotide. The gap is filled by the DNA epsilon by the 5 prime to 3 prime activity and the DNA ligase join the Okazaki fragment of the growing DNA strands. So this is all what we are going to show the lagging strand synthesis and as well as so in the lagging strand synthesis you are going to have the synthesis of the RNA, you are going to have the synthesis of unwinding the DNA and so on.

Combined activity of DNA pol delta and DNA pol epsilon. So looping of the lagging strand allow a combined DNA polymerase delta and DNA polymerase epsilon asymmetric dimer to assemble and elongate both leading and lagging strand in the same overall direction of the fork movement. And then the last portion or the last step is the terminations. So when the replication fork meets each other then the termination occur. It will result in the formation of the two duplex DNA. Even though the replication is terminated 5 prime end of the telomeric part of the new silicidase DNA molecule found to have shorter DNA strand than the template strand.

This shortage is corrected by the action of an enzyme the only actual replication is completed the enzyme is called as the telomerases. So when the replication is happening it is going to start from the center and one of the fork will go in this direction the other fork will go in this direction and that is how when they will meet with each other or they will meet the fork of the other the general replication is going to stop. But at these corners that these corners what will happen is that a synthesis is not going to be complete because the last primer what you are going to use is not going to allow the synthesis of the last portion of the DNA and these portion are actually going to be synthesized by an enzyme which is called as telomerase. And once the until the telomerase does not come

and synthesize the telomeres it is very difficult to say that the DNA replication is complete because if that does not happen then this portion is going to be lost and if it is keep losing then there will be a shortening of the genomic DNA. Telomerases are the enzyme which actually synthesize the telomeres or which actually are going to since complete the synthesis of the telomeres.

So during eukaryotic replications telomerase play a crucial role in ensuring the accurate replication of the linear chromosome. So the genetic cell uses a semi-conserved application to replicate this DNA and this process poses a challenge at the end of the linear chromosome. Telomerases serve to overcome this challenge and maintain the integrity of the genetic information. So at the corner of the chromosome these corners are not going to be replicated because there will be a because it is going to be a problem. So a short stretch of the 5 to 8 base pair and the tandem repeats are and the G series nucleotide sequences are actually going to be a problem for telomerase.

So telomerase is actually going to fill these gaps. Now what is the end replication problem? So linear genome including those of the several viruses as well as the chromosome of eukaryotic cell force a special problem completion of replication of the lagging strand. Contamination of the RNA primer from the 5 end of the linear molecule would leave a gap known as the primer gap and this primer gap cannot be filled by action of DNA polymerase because of the absence of the primer terminus to the extent and if the DNA does not get replicated the chromosome will shorten a bit with each round of replication and this problem has been solved by an enzyme which is called as the telomerases. So telomerases also known as the RNA dependent DNA polymerase and it is a ribonucleotide containing the RNA component having repeat of 9 to 13 nucleotides long. This RNA component serves as a template for the synthesis of the telomeric repeat at the parental DNA ends. So telomerases uses at the 3 prime end of the parental DNA strand as a primer and its template self DNA component.

So it is going to use the self DNA complaint as a template and it has a 5 prime to 3 prime RNA dependent DNA polymerase activity due to which it add the successive telomeric repeats to the parental DNA strand at its 3 prime end. This means the enzyme itself is going to have the RNA component and that RNA component is actually going to serve as a template and that is how it is actually going to synthesize the DNA complementary to that particular sequence and it is going to add that repeat on multiple occasions and that is how it is actually going to fill the last gap of that particular DNA. So regeneration of the telomerases, telomeric DNA consists of the simple tandem repeated sequence at the 5 prime end which are in the human for example you have the A GGG TT in higher plants you have the A GGG TT TT algae you have the A GGG TT TT protozoan you have the GGGG and TT TT and the yeast you are going to have the

GGG and T. So these are the repeat sequences what are being present into the telomeric regions and the telomerase is going to add these repeats on multiple occasions. Telomerase uses its RNA component as a template and parental DNA as a primer and then by its RNA dependent DNA polymerase activity it repeatedly add the telomeric sequences to the 3 prime end of the parental DNA and then it is released.

At last the RNA primers of telomerase is bound near the leading lagging strand and it is extended by the DNA polymerase. Thus the lagging strand synthesis is completed. So this is what exactly what is going to happen. So for example you once you have the this kind of situation where this portion is actually going to be need to be synthesized. So telomerase is actually going to bind to the 3 prime end of the telomer and that is complementary to the telomeric RNA and that is how it is actually going to extend.

So bases are added using RNA as a template. So it is going to have this particular type of RNA what is been already been present inside the telomerase and it is going to synthesize this. So it is that is how it is actually going to synthesize this sequence and you remember this is actually having a U. So instead of U you are going to have the A and that is how it is actually going to be keep synthesizing this and utilizing this DNA polymerase will actually going to bind a complementary DNA lagging strand and it is actually going to synthesize this strand and that is how you are going to have the completion of the synthesis of the telomerases. Elizabeth Blackburn and her colleague have provided the answer to fill up the gaps with the help of the telomerases known as the modified reverse transcriptase or telomer transferase. So that the genes are end are conserved in human the RNA template consist of AAU CCC repeats.

Examples of repetitive sequence which varies among species for example, Tetrahyena is going to have the AA triple C and Oxyrhittica you are going to have the AAA triple C. Now this is the mechanism through which sorry so this is the mechanism through which the telomerase is going to fill the gaps you are going to have 3 prime end of the lagging strand base pair with the unique region of the telomerase associated RNA. The telomerase catalytic site at the deoxynucleotide using the RNA molecule as a template and the telomerase then translocate to the new 3 prime end by pairing with the RNA template and it continue with the reverse transcription. DNA polymerase uses the newly made primer for the synthesis of DNA to fill the remaining gap.

The primer is the removed and nicked between the fragment is sealed by the DNA ligase. So this is all about the DNA applications in the prokaryotic and as well as the eukaryotic system. What we have discussed so far we have discussed about the DNA application in the eukaryotic system and how it is different from the prokaryotic system in terms of the machinery and in terms of the processivity and in terms of the processes

and you can you might have now realized that it is very complicated in the eukaryotic system compared to the prokaryotic system. So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss some more aspects related to the DNA applications. Thank you. .

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 05
Replication
Lecture-24 Mutagenesis and repair mechanism

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati and what we were discussing we were discussing about the different aspects of the molecular biology in this particular MOOCs course. So far what we have discussed in this particular module we have discussed about the DNA applications in prokaryotes followed by we have also discussed about the different steps of the DNA applications in the prokaryotic as well as the eukaryotic system. So we have discussed about the semi conservative mode of the DNA applications followed by we have also discussed about the DNA replication machinery in the prokaryotic and as well as the eukaryotic system. And in the last lecture we have also discussed about the importance of the telomerases and the telomerase enzymes and how it is actually helping to fill the gaps within the at the termini of the chromosomes. Now when the DNA replication is happening or even the cell is actually been continuously been exposed to the different types of the mutagenic molecules or the different types of mutations happens within the cell because when the cell is replicating it has a proofread activity it has utilizes that proofread activity to correct the sequences.

But even then there are spontaneous mutations which are going to happen into the DNA and these mutations are actually going to be detected by a well established machinery and that is how it is actually going to participate and it is going to correct those mistakes. You know that if these mistakes mutations are not been corrected they are actually going to lead to the accumulation of mutations and ultimately it may actually cause the development of the different types of the cancers. So DNA damage as the name suggests it could be occurs into a DNA due to the error into the DNA replications it could happen because it caused the spontaneous lesions or it could be because of the transposable DNA or it could be because of the physical mutations or the chemical mutations. Physical mutagens there are several examples of the physical mutagen and there are several samples of mutagen which we are going to discuss and then errors in the DNA replication is also very very important in terms of the DNA damage.

So what is mean by the DNA damage? It is known that the genome is not a static entity right genome is present inside the nucleus or inside the cell right in the case of prokaryotic system and it is whenever you are actually getting exposed to a particular type of physical or chemical environment it is actually getting exposed as well. So and

you know that the importance of the genome right is actually hereditary material so it is actually going to follow the information to the next generation and that is not only common so it is important that the genome should be intact and it should not be get damaged. So hence it is highly subjected to a variety of heritable changes. A sudden change in the sequence of an organism genome that give rise to the alternate form of any gene is called as the mutations as a result the DNA get damaged. These mutations are mostly recessive and lethal mutations are random can occur anytime in any of the cell of an organism it is not like that if the mutation occur once the same mutation cannot occur again mutation is recurrent can cause again and again right.

So the question comes how does the DNA damages happens? On the molecular basis DNA can be damaged in two ways some mutation may cause spontaneously some mutations are induced by the mutagens. Spontaneous damage occurs without the treatment of the organism with an exogenous mutagens it is mainly due to an error in replications spontaneous lesions and the transposable elements. Whereas the induced mutations arises when the mutagen react with the parent DNA which causes the structural alternation in base pairing. What is mean by the mutagens? Any agent which causes the frequency of the mutations is called as which increases the frequency of mutation is called as the mutagens. Now as far as the DNA damage is concerned you can have the two main category one is called as spontaneous damage the other one is called as induced mutations.

So within this category you have the multiple components and these components are important for causing the DNA damage. So agents causing the DNA damage you have these agents which are causing the spontaneous mutations and you also have the agent which are causing the induced mutations. Within the spontaneous mutations you can have the errors in the DNA applications you can have the tautomeric shifts and spontaneous lesions within the spontaneous lesion you are going to have the deaminations, depredations and deprediminations and the oxidative cleavage. Whereas the transpositions is another phenomena through which it is actually going to cause the spontaneous mutations. Whereas in the case of induced mutations you are going to have the two different categories which are can cause the induced mutations.

One is called as the chemical mutagens otherwise called as the physical mutagens. Within the chemical mutagens you are going to have the different types of agents which are going to cause the base analogs. So you are going to have the chemical agents which are going to function as a base analogs and that is how they are actually going to destroy the DNA. Then you can also have the deanimating agents you are going to have alkalating agents. Most of the anti-cancer drugs or majority of the anti-cancer drugs are alkalating agents and then you are also going to have the intercalating agents because

they all are actually going to do interfering into the replications and interfering into the repair mechanism and that is how they are actually going to cause the mutations and that mutation will perpetuate from the generation to generation.

Then you are also going to have the physical mutagens. Physical mutagens means the physical parameters what you are going to use. So in that case you are going to have the UV radiations, you are going to have the ionization radiations and you are also going to have heat. So these are the physical mutations which are physical parameters which are actually going to cause the mutations. Now let us first start with the spontaneous mutations and then we will discuss about the induced mutations.

So spontaneous mutations number one category is the error into the DNA applications. So tautomeric shifts. For example so shift of a proton from its nitrogenous base form its rare form and that is called as the tautomeric shifts. The stable keto form of the thymine and adenine and the amino form of the adenine and cytosine undergoes tautomeric shift to form the unstable enols and imines respectively. For example you have the stable form of the adenine and guanine that is the amine form and once it goes from the tautomeric shifts it is actually going to form the immuno forms and these immuno forms are very rare but they are actually going to be less unstable.

These forms are very short lifespans if they are incorporated into nascent DNA they may result into the mutations. And these bases are present in their unstable enol or the immuno state they tend to form the AC and GT base pairing. This AC and GT base pairing is not allowed and that is how it is actually going to destroyed or it is actually just going to distort the DNA structures and because of that for example this is C it is actually going to form the interaction with the A form. So A is always making a base pairing with T in the DNA but in this case A is actually going to make a pair with C and that is actually going to cause the alteration into the sequences and that is how it is actually going to go into the next generation. So imagine that you have a C so instead of if you have a C into the template ideally it should be G into the replicated DNA.

But since this kind of mutation is happening or this kind of tautomeric shift is happening what will happen is that C is actually going to give rise to a synthesis of A. So instead of A instead of G you are going to have A into the DNA sequences and that is how it is actually going to cause the mutations. The net effect of such event and the subsequent replication required to segregate and mismatch pair that is the AT to GC or GC to AT base pairing substitutions and that is how it is actually going to cause the mutations. Mutations via the tautomeric shift in the basis of DNA in the examples a guanine undergoes a tautomeric shift to a rare in all form at the time of replications. In its in all form it is paired with the thymine.

So this is what exactly is showing here that you are actually replicating a sequence where you have the all sort of nucleotides like G, C, T and all that and during the replication if the G is actually getting converted or tautomeric getting shift into a in all form then G is actually going to start making a pair with T instead of G is going to start making pair C. And because it has altered the base pairing information it is actually going to allow the formation of the instead of GC it is actually going to give you the AT because it is actually going to give you the T instead of G and as a result it is going to propagate. So once the DNA replication is going to happen of this particular sequence it is actually going to give you the in the wild type it should be G is actually going to give you the C but in this particular type of second generation instead of G and C it is actually going to AT. So this is what exactly happened when you are going to have the tautomeric shifts. So during the next generation the joining shift back to its more stable keto form thymine incorporated opposite the in all form of the guanine direct the incorporation of the adine during the subsequent replication.

The net result is that GC is going to be AT mutations. So G is making a base pairing with T so as a result in the first generation the T is going to be synthesized and then in the second generation G is again reverting back to the in all form into the into the keto form. So because of that it is actually going to start again synthesizing the G to C but here instead of T it is actually going to since you have the T in the daughter it is actually going to start synthesizing the A and as a result what will happen is that wherever you have in the G it is going to be replaced by A. So it is actually going to have the G to A mutations and if you have G to A mutations then it is actually going to change many things it is going to change the amino acid what is corresponding to the G to A substitution and so on. And then you have the substitution mutations so and during replication itself to the substitution mutation or the frame shift mutations.

So substitution of one base pair into another is called as a substitution mutation the swapping of the base pair may be a transition mutation or the transversal mutations. So what is the transition mutations? Substitution of one pyrimidine by another pyrimidine or one purine by another purine is called as the transition mutations. For example G to C is going to be replaced by A and T and vice versa. So in that case it is actually going to cause the substitution mutations. Then you have the transversion mutations so purine is replaced by the pyrimidine or the pyrimidine is replaced by the purine.

So subsequent transversion is transition for example G to A or A to G or C to T or T to C. These are the if the G is replaced by A then it is going to be called as transition mutations or A is replaced by G then it is also going to be called because the pyrimidine is replaced by the pyrimidine and purine is replaced by the purine. But if it is the

pyrimidine which is replaced by the purine then it is going to be called as transversal. So both of these kind of substitutions are actually going to be very very problematic because it is going to overall change the amino acids and it is going to cause the mutations into the subsequent gene products. Then you have the frame shift mutations.

So sometime it may happen that during the application some extra nucleotide may get inserted or may get copied. So you can actually have the either the addition of the extra nucleotides or you can actually have the disappearance or the deletion of some of the mutations and in both of these cases there it is actually going to cause the frame shift mutation. This frame information anyway you will be able to understand when we are going to discuss about the translation because you know that the protein is going to be synthesized in the form of the codons and codon is made up of the 3 nucleotides. So protein is been synthesized in the form of the codon. So for example this is

the	codon	ATG.
-----	-------	------

So this 3 nucleotides are going to be read together by the anti codon what is present onto the tRNA. But now imagine that if I add one more A into this so what will happen is it is going to be a this sequence. Now if I go by the 3 triplicates then it is actually going to be like this. This means earlier the codon was ATG now the codon is going to be ATA and codon for ATG is going to be X amino acid whereas codon for ATA could be Y amino acid and that is how there could be a change in the amino acid what is going to be incorporated into the protein. The another example is that for example if I have the

A	here	and	if	I	or	C	for	example.
---	------	-----	----	---	----	---	-----	----------

So if I have a sequence like ATG C A and if I remove actually if I remove for example if I remove this actually. So if I there will be a deletion then what will happen is it is going to have this. So earlier you are having this as a codon now you are going to have this as a codon. So this also is going to be cause of deletion. So this is going to happen because of the deletion this is going to happen because of the So if this happen in the exorheic region of DNA that may mainly change the translational reading frame has results in the production of a non functional protein which can be observed in a phenotypic characteristics.

So it is not only going to change the codon for this it is actually going to change the codon for subsequent generation also. For example it is going to change the frame shift rotation. So it is actually going to change the codon for example earlier you are having this. So this is the first codon this is the second codon. Now when you have removal of this G this G remove if you remove this T then you are going to have the first codon as AGC and the second codon would be ATT whereas earlier it was ATG and CAT.

So this was the codon. So this is not only going to change the frame shift only for first codon but also for all the codons and because of that it is actually going to cause a significant change into the amino acid composition of the product and as a result it is actually going to change produce the protein which may not be functional and which may actually cause the problem to the cell. For example this is the example what is being given you have this is the standard frames what is present and when you have the mutations or messenger RNA what is going to be formed it is actually going to form but if you have the addition of T base pair then it is actually going to change. It is going to change the frames and that is how it is actually going to form the different mutations or if suppose you have the deletion then also you are going to change the things. So these are the some of the things which are very very important and the frame shift mutation is very very significant problems and as a result this shift to change in a translational reading frame hence name as the frame shift mutations. Then we have the second thing is spontaneous lesions naturally occurring damage to the DNA is called as the spontaneous lesions most common lesions are deaminations and depurinations.

So deaminations the loss of exocyclic amino acid group from the cytosine and adenine and the guanine due to change in the pH in temperature is spontaneously resulting in the formation of the uracil, hypoxanthine, xanthine and thymine respectively. So deamination is going to change the amino acid for example if there will be a deamination from these of the from of these nucleotides then it is actually going to be get converted into uracil, hypoxanthine, xanthine and thymine and that is actually going to result into the change in the structure or change in the nucleotide sequences. Then depurinations and depurimidations loss of purines and pyrimidine by the breakdown of the glycosidic bond nucleotide from the DNA due to damage due to the change in the pH and then you also going to have the oxidative damage, damage in the DNA due to the reactive species spontaneously radicals like peroxide, hydrogen peroxide and hydroxyl radicals attacks on the DNA probe product reduce variety of products attack at several leads to the fragments base loss and the strand streaks. So all of these things can also lead to the frame shift mutations and the spontaneous mutations as well. So these are the some of the example if you have the deaminations the cytosine is going to be get converted into uracil if there will be a deamination into the iodine it is actually going to be get converted into the hypoxanthine and then if you have the 5 methyl cytosine and if there will be a deamination it is going to be get converted into the thymine and if there will be a deamination into the guanine then it is actually going to be get converted into xanthine and some of these are non natural nucleotides so they will actually going to cause the significant mutations into the DNA structures.

And then we have the transpositions transposon is a DNA segment that have the capacity to insert itself at any location in the genome without having any relation to the

target sequence consequently it cause loss of gene fraction or genes inappropriate over expressions. So transposition is a very important and the very significant topic that you should actually study and transposons are also called as the jumping genes and they are actually changing its position from the one locus of the genome to the another locus and as a result they are actually causing the different types of artifacts and different types of problem into the genes. So they are actually going to cause the gene loss of gene functions or genes appropriate over expressions. So transposon is a very important and the big topic so that you can actually be if you are interested you can be able to study this from any of the standard another molecular biology books. And then we have the induced mutations.

So mutagens results into the induced mutations mutagens can be classified as a physical mutagens and the chemical mutagens. So these are the physical mutagens where you have the UV radiations, ionizing radiations and the heat. So UV radiations UV radiation is a potent physical agent that cause a number of photo product in the DNA. UV radiation of the wavelength 260 nanometer induced dimerization of the pyrimidine nucleotides bases especially thymine resulting in the formation of the cyclobutyl dimers. So this is what exactly happened when you have the two nucleotide they will be getting exposed to the UV they are actually going to get dimerized and they are actually going to form the cyclobutene rings.

Adjacent pyrimidines are covalently linked to the formation of four member ring structure and this structure is called as the pyrimidine dimers. Then ionizing radiation mainly causes the DNA strand breakage and then you also have the heat stimulate the wave induced cleavage of the N glycosidic bond which result in the apirenic or apiremitinic side or the baseless sides. So in this particular thing what will happen is that there will be a cleavage of the glycosidic bond and you know that the glycosidic bond is holding the sugar to the base and because of that there will be no base what is present on to these particular nucleotides. Then we have the chemical mutagens so in the chemical mutagens you can have the four categories the base analogs, de-animating agents, alkylating agents and the intercalating agents. Base analogs are certain bases which are not present in the DNA normally but resembles to the normal nitrogenous bases that can incorporate during the DNA synthesis.

For example the 5 bromo uracil which was base analog to the thiamine and the 2 amino purine which is analog of the adenine. So these are the base analogs are going to induce the mutations because they are going to mimic the natural bases but they are not the natural and they will be get incorporated into the DNA during the DNA synthesis. Then we have the de-animating agents which causes the point mutation by removal of the amino group from the nitrogenous bases. Nitrous acid de-aminate the adenine, cytosine

and guanine. Then bisulfite de-amine only cytosine and the de-amination of adenine give rise to the hypoxanthine which pair with C instead of T and de-amination of cytosine give rise to uracil which pairs with A instead of G and this actually is going to cause the problem into the first generation and as well as into subsequent generation because of the C to T mutations and A to G mutations.

Then we have the alkylating agents so some example of alkylating agents are like ethylene methane sulfone or the nitrogen mustard and the dimethyl nitrosamine. These agents actually add the alkyl group to the certain position in the nucleotide and these alkylating agents are not actually going to cause the mutation and that is why it is actually going to kill the cells. Many of the anti-cancer drugs are also alkylating agents so they alkylate the DNA and that is how they are actually activating the machinery to you know to kill these cells. And then we have intercalating agents usually associated with the simple single mutation pair insertion or deletions. Intercalating agents are flat molecule that flip between the base pair in the double helix resulting into the unwinding of the DNA helix and therefore increasing distance between the adjacent base pairs.

Some examples are proflavine, acridine orange, ethidium bromide and the ICR compounds. Intercalating agent is also very, very problematic because they were going to interfere into the DNA synthesis and they are also going to change the normal DNA structures and that is how they are actually going to cause the mutations. So these are the one of the effects by which the so for example, if this is the DNA, how the base analog is actually causing the problem. So for example, this is the normal DNA and you have the A which is and in the template in the corresponding template you are going to have T. Now, if there will be a base mutation, so example the 5 bromo uracil is going to be present inside of A, then the BU is actually going to so if there will be a BU undergoes tautomeric shift then BU is actually going to get converted into this and that is how it is actually going to have the interaction, it is going to have the affinity for the C rather than G because this is going to have the affinity for G rather than C, rather than A because A U should have an affinity for A, but it does not have the affinity for A, it has the affinity for G.

So, instead of so what will happen is when the replication occurs in the wild type the A is going to form and A is going to be synthesized instead of in front of T, but in this one instead of this you are going to have the G. Now, once you do the another round of applications A this wild type will not actually have any problem because it is going to still have the A to T, but here now the G is first mutated and now the second strand is also going to be mutated and that is how it is actually going to have the C. So, this is actually going to be a mutated DNA what is going to be formed and this mutated DNA if you do the subsequent mutations subsequent replications will proliferate and it will

continue into the subsequent cells actually or notodotus cells. Similarly you may have the deaminating agents for example, adenine when it is going to be attached to the nitrous acid it is going to get converted into hypoxanthine or the cytosine when it is going to get converted into AT. So, once you have the adenine to the hypoxanthine, hypoxanthine will not go to form the interaction with T instead it is actually going to have the interaction with C and because of this the AT is actually going to be replaced by GC exactly with the same mechanism that in the first generation the A will actually go to recognize the T, but in the subsequent generation when the A will get converted as hypoxanthine the daughter strand is going to have the cytosine and cytosine is actually going to incorporate the C.

So, that is what exactly going to happen in the subsequent generation in with the deaminating agents and then you also going to have the alkylating agent for example, guanine is going to get converted into the ethylene guanine and ethylene guanine is going to have the affinity for thymine instead of the interaction for cytosine and because of that there will be a reverse. In this case the AT is being replaced by GC in this case the GC is going to replace by AT and the same is true for the thymine when it is getting alkylated it is going to form the ethylene thymine and ethylene thymine have a interaction or going to make the base pairing with the guanine rather than the adenine and as a result the TA is going to replace by the CG. So, these are the some of the mechanism through which the DNA is going to be mutated. Now how the DNA if the DNA got damaged whether it is going to be damaged by the spontaneous mutations or it is going to be damaged by the induced mutations you are supposed to repair these changes. But supposed to have the machinery to detect and then if possible you can actually be able to have the repairing of these things.

So systems for repairing the numerous unintended lesion that frequently occurs in DNA are also required for maintaining the genetic stability in addition to the extremely precise DNA replication mechanism. The vast majority of these unintentional DNA replications are temporary because the DNA repair a group of connected system immediately correct them. Without repair processes a genome cannot maintain its essential biological functions. These DNA repair mechanism fall into the two broad categories direct reversal of the chemical process that cause the DNA damage and the damage bases are removed and then replaced with the freshly synthesized DNA. So you have the two choices of you know reversing these damages.

First you can actually be able to reverse the process you can actually reverse the reaction for example if you have the alkylating agents you can actually reverse the alkylating agent and you know reverse these reactions so that the guanine is going to be converted into back to the adenine and so on. Another thing is that you can just replace

these damaged base and replace it with the normal base and that is also going to be another base. So additional mechanism have developed to help the cell deal with the damage where the DNA repair fails. DNA repair mechanisms you can actually have the single stand break repair mechanism or the double standard break mechanism. In the single stand break mechanism you can have the direct reversal, excisions, mismatched repairs and within the excision repairs you can have the nucleotide excision repairs or the base excision repairs.

Whereas in the double strand breaks you can have the homologous recombination or the non homologous DNA recombination. So single stand breaks you can have the direct reversal so directly acting on the damaged nucleotide converting back to the original structures. So pyrimidine dimers are usually repaired by a light dependent direct reversal process called as the photo reactivations. DNA photolyase enzyme participate in the mechanism which is found into the E. coli these enzyme get activated with the wavelength of the 300 nanometer and the 500 nanometer the enzyme binds to the pyrimidine dimer and convert back to the original monomeric and nucleotides.

So this is what exactly happened when you have the thymine dimers which are going to be formed and when you are exposing this particular DNA with a visible light it is actually going to have the photolyases of this particular strands where you are actually having the nucleotide dimer which is being formed. And then with the help of the enzyme which is called as the DNA photolyase it is actually going to replaced and it is going to remove these nucleotides and it is going to form the normal DNA. So these are going to reversed and that is how it is actually going to recover the DNA. Then we also have the excision repairs so excision repairs it involves the excision of the damaged segment of DNA followed by the re-synthesis of the current nucleotide sequence by an enzyme which is called as DNA polymerase. So base excision repairs you can actually have the removal of the damaged base used to repair minor damages like the alkylation, deamination which are consequence of exposure to the mutagens.

It is initiated by enzyme which is called as DNA glycosylase and the DNA glycosylase cleaved at glycosidic linkage and detaching the altered base as a result an a-pyrenic or a-pyramidinic sites are generated. Final ligation of the nucleotide take place by the DNA polymerase and the DNA ligase. So these are the this is the mechanism where you are actually having the UDG role of the UDG uracilidin and glycosylase. So if you have the deaminated cytosine which is actually going to be uracil and that is actually going to first form the a-pyrenic sites and then this a-pyrenic sites are actually going to refilled back and that is how you are going to have the repaired DNA at the end of this particular repairing mechanisms.

Then you can also have the nucleotide excision repairs. So similar to base excision repair it acts on a more substantially damage area on the DNA it include the following steps. So damaged DNA first you are going to have the damaged DNA the breakage of the phosphodiester bond on the either side of the damaged portion. So for example this is the damaged DNA so on the both side you are going to have the breakage of the phosphodiester bond and then you are going to have the excision leaves gap. So you are going to have a gap which is going to be created and this gap is going to be filled by the DNA polymerase and ligase seal the breaks. The best studies mechanism for the DNA replays is the UVR system into the E.

coli a key enzyme is involved called ABC exonuclease which is made up of the three subunit UVR A, UVR B and UVR C genes and ABC XC nucleus binds to the damaged site on the DNA and cut the phosphodiester bond on the both 5 prime and 3 prime. UVR D is acting as a helix case and helps to unwind the DNA at the site of the cut the gap is filled by the DNA polymerase one and sealed by the ligase. And then we have the third mechanism the third mechanism is called as a mismatch repair. It detects the mismatch that occurs during the renin applications. The mismatch repair occurs in daughter strand and it highly prone to the mismatch during the replications.

How to distinguish between the parent strand and the daughter strands? Immediately after the replication parent strand will be containing methyl group whereas the daughter strand await for the introduction of the methyl group. In this way the two strand can be distinguished and this is the best time period for the repair mechanism to scan the lesion like the correction. So this is very important that see the damage will only going to occur into the daughter cells rather than the parent cell. So parent cell you are just remove going to remain as intact in the daughter cells you are going to have these kind of repair mechanisms. And how you are going to recognize the daughter cell the daughter DNA the daughter DNA is going to be unmethylated because the methylation still not been done for all the adenine groups.

How the methylation occurs? In E. coli the methylation is executed by an enzyme which is called as the DNA adenine methylase which converts the adenosine to the 6 methyl adenosine in the sequence sequences like GATC and the DNA cytosine methylase which converts cytosine into the 5 methyl cytosine in the 5 prime sequences. It should be noted that these methylations are not mutagenic the modified version have the same base property as the unmodified one. And then E. coli the mechanism is completed by the mut protein there is a involvement of 3 mut proteins the mutH mutL and mutS mutH and mutL recognize the sequence GATC and mismatch respectively. And then we also have the SOS response so sometime a DNA damage is so dangerous that is stimulate the cells to produce the DNA repair enzyme that allows an immediate

reaction to the specific DNA damage.

SOS response in bacteria is the most well researched illustrations when the bacterial chromosome is severely damaged the SOS response genes are activated many of these genes are involved in a repair and mutagenesis. RecA protein and LexA repressors are the main player in the SOS response. LexA is a repressor that binds to the 20 base pair segment of DNA known as SOS box to prevent the activation of the SOS response genes. As a result the LexA governs the transcription of the rest each SOS response genes. So this is just a mechanism through which the SOS response is happening. You are going to have the SOS box where the LexA gene is going to bind and that is how it is going to start the SOS response gene transcription.

And there will be no transcription if there is no damage to the chromosomal DNA. But if there will be a damage to the chromosomal DNA then there is RecA you are going to have the recruitment of RecA and that is how it is actually going to cause the cleavage of LexA and once there will be a cleavage of LexA then this inhibition or the attenuation is going to be removed and that is how the SOS box is free and that is how it is actually going to cause the production of the SOS response gene and that is actually going to cause the bacterial chromosome damage repairing. And then we also have the double strand breaks so far what we have been discussing we were discussing about the single stranded break repair mechanisms. In the double standard break mechanisms you have the two mechanisms the homologous recombinations and the other kinds of mechanisms. So these are some of the steps what you are supposed to follow when you are going to have the homologous recombinations.

And these steps are very very important and we are going to discuss in detail of all these steps when we are going to discuss about the genome editing and other kinds of phenomena in a subsequent lectures. And then we are going to have the non homologous DNA repair and joining and all these we are actually going to discuss because the non homologous process repair double standard breaks in the DNA and this process is known as homologous because of the break DNA are directly ligated without any homologous template. This was pointed by the Maury and Haber in the year of 1996 and double standard breaks are recognized by a protein which is called as Ku 70 by 80 and the Ku 70 by 80 recruit the DNA Pks or the kinases and then recruit the Artemis which remove the damaged ends. So in conclusion DNA damage is inevitable due to the various factors cells have evolved DNA repair mechanism to counteract the DNA damage. Types of DNA damage includes the breaks, base modification and cross links repair pathway like the NER, MMR and HR fix the different type of DNA damages.

Accumulated unrepaired DNA damage can lead to the mutation and disease including the cancers. Nerve repair mechanism contribute to the cancer development. Some cancer therapies exploit the DNA repair mechanisms and that is how they can actually cause the mutations into the cancer cells and since you are going to have several mutations the cancer cell will have no option but to die actually. Aging is linked to the declining DNA repair efficiency the environmental factor can overwhelm the DNA repair mechanisms and ongoing research are seek to improve the understanding and develop the therapies. So there are many mechanism the DNA once the DNA is damaged there will be many mechanism through which the DNA is been repaired into the eukaryotic cell or the prokaryotic cell.

So this is all about the DNA damage and repair mechanisms and we have discussed various mutagenic molecules which are actually causing the induced mutations. We have also discussed about the spontaneous mutations and how the spontaneous mutation is happening because of the error prone because of the DNA replications and the other phenomena. So with this I would like to conclude my lecture here and subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 06
Transcription
Lecture-25 Transcription in Prokaryotic System

Hello everyone, this is Dr. Vishal Trivedi from the department of biosciences and bioengineering IIT Guwahati and what we were discussing we were discussing about the different aspects related to the central dogma of molecular biology. So in the previous module we have discussed about the replications and where we have discussed about the replication in the prokaryotes and replication in the eukaryotes. Apart from that we have also discussed about how the replication is playing a pivotal role in making the DNA damage and repairs and if you recall in the previous lecture we have also discussed about the what are the different other processes occurring in the central dogma of molecular biology. So what we have discussed that the central dogma of molecular biology is a very concentrated and very regulated process in which the DNA is going to be multiplied so that the DNA can be divided between the mother cell and the daughter cell and on the other hand the DNA is going to give rise to the RNA and that RNA is going to be responsible for the production of the proteins. So the production of or the synthesis of the RNA from the DNA is been catalyzed by the RNA polymerase and this process is known as transcription.

So in today's lecture we are going to discuss about the transcription and how the transcription is actually going to be done in the prokaryotic as well as the eukaryotic system. So in today's lecture we will discuss about the prokaryotic genes eukaryotic genes and how the prokaryotic transcription in the prokaryotes is been done versus the prokaryotic transcription in the eukaryotes. So when we say the transcription means that there will be a synthesis of the RNA from the DNA and this process is called as the transcription and this process is been catalyzed by a enzyme which is called as the RNA polymerase. And RNA polymerase is a very multimeric protein complex and it has the sub multiple subunits both in the prokaryote and as well as the eukaryotic cells and the RNA polymerase is a very very important enzyme for giving the DNA from the RNA.

Now before we getting into the details of the transcription it is important for us to understand the about the gene and how the gene is been different from the prokaryotes versus the eukaryotes. So what is transcription? So as we already discussed the transcription is the synthesis of the RNA from the DNA. So every cell contains three different types of RNA that is the transfer RNA or the tRNA, ribosomal RNA or the rRNA and the messenger RNA or the mRNA. Synthesis of RNA from the DNA template

with the help of the DNA dependent RNA polymerase is known as transcription. It occurs unidirectionally in which the chain is synthesized in the direction of 5 prime to 3 prime.

The segment which is transcribed from the DNA is known as the transcription unit. So the DNA the synthesis of the DNA to RNA is called as transcription and we have the three different types of RNAs. We have the ribosomal RNA, we have the ribosomal RNA, we have the tRNA and we have the messenger RNA. And if you recall when we were talking of when we were discussing about the biomolecules we have discussed about the structure of the messenger RNA and how you can be able to purify the messenger RNA from the cell with the help of the two different methods we have discussed about the trizole method and we have also discussed about the affinity purification as well. So the role of these RNA molecules are different right messenger RNA is actually being carried the message right.

So and it is required for providing the message in what sequence the amino acids are actually going to be attached. The tRNA is actually going to be carry the amino acid. You know that the proteins are made up of the amino acid that also we have discussed in the previous module. So it will actually going to carry the amino acid on one side and on the other side it also going to carry the information so that it will actually going to recognize carry the information to recognize the messenger RNA right. So it is actually going to carry it is going to recognize it is going to carry the amino acid and as well as it is actually going to recognize the messenger RNA and then the ribosomal RNA it is actually going to synthesize the protein.

So it is actually going to form the ribosomes and that is actually going to synthesize the protein by forming a peptide bond between the amino acid which is going to be carried by the transfer RNA. So basically all these RNA molecules are going to have one or other functions and they are actually going to be formed from the DNA which is by a process which is called as the transcription. So we have separate genes for synthesizing the ribosomal RNA, tRNA and messenger RNAs and the segment which is transcribed from the DNA is known as the transcriptional unit. So in a eukaryotes it is actually the monosystronic transcriptional unit which occurs to code for single polypeptide whereas in the prokaryote it is actually going to be polycystronic. It means it is actually going to transcribe and it is actually going to give you the more than one polypeptide.

So in a transcriptional unit what you have you have a promoter right this is the promoter then you are going to have the coding sequence and then you are also going to have the terminators and both are all these three segments have their specific role that promoter is actually going to provide a docking site for the RNA polymerase to initiate and in that

started whereas the RNA coding sequence is actually the coding sequence which actually is going to give you the RNA. It could be messenger RNA, it could be transfer RNA or it could be ribosomal RNA and once the synthesis is over then they are actually going to be a sequence which are actually going to bring the ending of these sequences. So promoter is actually going to be us going to help in the starting of the transcription replication transcription the coding sequence is actually going to allow the elongation and that is how the elongation will continue and the termination sequences are actually going to have the termination. So it is going to help in the termination and we are going to see the features of all of these components of the transcriptional units. So transcriptional unit as I said you know it is going to be having the promoters coding sequence and the terminators.

The promoter is going to be a starting point, it is going to be for elongation and this is for the termination. So what is the start point? It is the first base pair from where the transcription start and it is called as the start site. RNA polymerase from the moves from the start point along with the template synthesize the RNA up to the termination sequence and you have the upstream and as well as the downstream sequences. So upstream it is a non template nucleotide in the 5 prime end or the minus direction which is sequence before the start point. So this is actually the start point and before the start point you are going to have the promoter so that is going to be upstream sequences.

Then you also going to have a downstream sequences so it is a nucleotide in the 3 prime end or the plus direction so this is going to be the downstream directions and it is actually going to be a sequence after the start point. DNA is a double helix structure so during transcription only the one strand is transcribed so that the transcriptional sequence is identical with the one strand of the DNA known as the coding or the census strand and the other complementary strand is known as the template or the antisense strand. You know that transcription is going to occur in the direction of 5 prime to 3 prime this means it is actually going to read the information from the 3 prime to 5 prime strand. So it is actually going to read the sequence from the 3 prime to 5 prime strand and that is why so if you see that if the RNA polymerase is going to sit here and it is actually going to run in this direction on this strand then it is actually going to synthesize RNA which actually going to have the 5 prime on this side and 3 prime on this side and that is why this sequence is actually going to be a non-coding strand. This is going to be called as non-coding strand because this is actually going to be providing the template whereas this sequence is actually going to be a coding strand.

Before we discuss about the prokaryotic transcription and the eukaryotic transcription we will actually going to see the difference between the eukaryotic and the prokaryotic transcription. So what is the difference between the prokaryotic and the eukaryotic

transcription? Remember that the prokaryotic transcription is or prokaryotic genes are polycistronic whereas eukaryotic genes are monocistronic. So that is very very and that actually brings the difference in terms of their transcription. So the first difference is that the prokaryotic transcription or the prokaryotic genes are polycistronic which means they are actually going to code for many polypeptides whereas the eukaryotic gene is going to be monocistronic. So it is actually going to code for single polypeptide.

It means you are going to have the multiple genes present within the prokaryotic transcription unit whereas you are going to have the single gene. Since there is no nucleus right the prokaryotic transcription occurs within the cytoplasm right because the DNA is also present in the cytoplasm whereas the eukaryotic transcription occurs inside the nucleus and within the nucleus you are going to have the synthesis of the messenger RNA, tRNA and ribosomal RNA. Number 3 because the transcription is occurring in the two different compartments the transcription is not coupled with the translation in the case of the eukaryotic transcription because so it is not going to have the coupled transcription and the translation because the transcriptional unit and the translational units are present in the separate quarter right. So they are actually present in the separate compartment whereas in this case it is actually going to have the coupled transcription and the translation which means it is actually going to have the transcription as soon as the RNA comes out right it is actually going to be recognized by the translational unit and then the translation and the transcription will continue at the same time. Number 4 single type of RNA polymerase required for the synthesis of all type of RNAs.

So it is going to have the single RNA molecule RNA polymerase molecule which is actually going to be utilized for production for the synthesis of all type of RNA whether it is messenger RNA, tRNA and ribosomal RNA. Whereas in the case of eukaryotic system you are going to have the three different types of RNA polymerase which is required for the synthesis of the all different types of RNAs. Then number 5 you are going to have there is no need for any transcriptional factor for the initiation. So because why it is so because the RNA polymerase is competent enough to start the initiate the transcription whereas in the case of the eukaryotic transcription the eukaryotic transcription requires the transcriptional factor for the initiation. So actually the transcriptional factor are actually going to recognize the promoters and then only the RNA polymerase will come and bind.

So that anyway you will understand when we are going to discuss about the transcription in the eukaryotic as well as the prokaryotic units. The number 6 you are going to have the RNA polymerase are made up of T5 subunits whereas the RNA polymerase are made up of T10 to 15 subunits. So RNA polymerase is big and complex in the case of eukaryotic system whereas the RNA polymerase is small and simple in the

case of eukaryotic prokaryotic system. Now let us see what is the machinery of the prokaryotic system. So we will first discuss about the transcription in the prokaryotes and then we also going to discuss about the transcription in eukaryotes.

Many of these steps or the many of the basic steps are going to be remain same between the prokaryotic transcription and the eukaryotic transcription. So that we are not going to repeat when we are going to discuss about the transcription in the prokaryotes. So transcription in prokaryotes. RNA polymerase in prokaryotes a single type of RNA polymerase is present which is responsible for the synthesis of all types of RNA. In bacterial RNA polymerase is a holoenzyme and it is a multi subunit protein which contains the five different subunits alpha, alpha, beta, beta and sigma actually and alpha is

the assembly of the core enzyme.

So you are going to have the alpha alpha as the assembly of the core enzyme. So alpha is actually going to be to form the core enzyme whereas the beta and beta prime are going to perform all the enzymatic and catalytic functions. So this is the beta and the beta prime which is actually going to perform the enzymatic and the catalytic activities and sigma is actually going to recognize the promoter sequences. So you are going to have the core enzyme which is going to be formed by the two alpha and beta and beta prime whereas the holoenzyme is going to contain the core enzyme plus the sigma factors. And these sigma factors are going to be different for the different types of genes.

So that is how you are actually having a very simple system where you are going to have the RNA polymerase made up of the five different types of subunits and all these five different subunits can be divided into two part. One is the core enzyme and the other is the sigma factor and within the core enzyme you are going to have the two alpha subunits and the beta and beta prime and this is being done only to conserve the energy because if you are synthesizing the RNA polymerase according to the different types of promoters or according to the different types of RNA species what you require then you are supposed to synthesize a large number of RNA polymerases and RNA polymerase is a big enzyme and a big protein. So to conserve the energy what bacteria has decided or the prokaryotic system has decided that I will actually going to have the core enzyme which will actually going to have all the activities so that it actually going to have the it will be able to read the DNA sequences it will be able to have the you know the synthesis activity and so on. But to recognize the genes to recognize the promoters we are actually going to have the sigma factors and that is how you can be able to have the single core enzyme associated with the multiple type of sigma factors and that is how you can actually be able to utilize the same for the multiple genes and that is how you can be able to conserve the energy. Now the second part is the promoter so prokaryotic promoters are simple compared to the eukaryotic promoters.

So they are simple so prokaryotic promoters typically consist of a 40 base pair region located near to the 5 prime end of the transcriptional start site promoter region consist of the two 6 pair consensus sequences called primbo box or the tata box and the minus 35 region. Primbo box is a 10 base pair upstream of start point and it is having a consensus sequence of TATAT whereas minus 35 region has the consensus sequence of TTGSEA. So this is actually going to be the promoter the typical promoter what is present in the prokaryotic system where you are going to have this is actually going to be called as the start site. So this is actually going to be if this is going to be the start site to the 5 prime end which means to the upstream of this you are going to have the minus 10 region and then you are going to have the minus 35 region and within the minus 35 region you are going to have the sequences which is going to be TTGSEA whereas in the minus 10 region you are going to have a sequence which is called as TATAT which means you are basically going to have a combination of these two and length or the distances between the two these two region is actually and the nucleotide what are present in these regions are actually going to decide whether the promoter is going to be strong promoter or it is actually going to be a weak promoter and depending upon that you are actually going to have the different types of you know the transcriptional activity of a particular gene is going to be different. This means the synthesis of the protein molecules are completely going to be governed by the strength of the promoter and that is how you can actually be able to modulate the expression as well as the production of a particular protein simply by modulating this because in a prokaryotic system remember that the transcription and the translation is going to be occurred simultaneously.

So as soon as the RNA species is going to be formed and it is going to be present in the cytosol or it is actually still be you know doing the transcription it will be available for the translational machine and it is going to start synthesis. So that is why you can actually have the better control over the protein production during the transcription itself and that anyway we are going to discuss and we are going to discuss about the control mechanism within the transcription and where we are going to discuss about the different types of operons. So here what we have is these are the three important or four important events what is going to occur in the transcription. So transcription in the prokaryote occurs in four stages one is number one is the template binding number two is the chain initiation number three is the chain elongation and a number four is termination. So the number one step is when the template is going to recognize this by the RNA polymerase and that is how it is actually going to initiate the synthesis of the RNA.

So the number one event is the binding of RNA polymerase to the template DNA and the chain initiation. So the DNA duplex should be opened so that the RNA pole can approach to the single standard DNA templates. Efficiency of the initiation is inversely

proportional to the melting temperature that is the TM and AT rich region has the lower TM because of the double bond hydrogen bonding and then the triple bond in GC region and thus it is more stable. Therefore the AT rich is good for melting of duplex and easy to create the open promoter complex than the GC rich region. So this region where the RNA polymerase is going to go and sit and then actually going to break the DNA or the unwind the DNA should be AT rich so that it should be easy for RNA polymerase to find the single standard DNA and that is how it is going to initiate.

So RNA polymerase has sigma factor so and you know that there are different types of sigma factors for the different types of genes. So you are going to RNA polymerase has sigma factor which recognize the promoter sequence at which the RNA polymerase holoenzyme binds and forms a complex which is known as the closed complex. In fact the sigma factor is released when the chain reaches nearly to the 10 base pair leaving the core enzyme for the elongation. So what happen is that suppose this is the promoter region right. If this is the promoter region, it is promoter region is actually going to be recognized by the sigma factor and as well as the sigma factor will go and bind then the RNA polymerase will come and it is actually going to bind the sigma factor and then this complex is going to unwind the DNA it is going to form the single standard DNA and that is how it is actually going to start the synthesis of the RNA.

So it is actually going to start running so RNA polymerase is actually going to start running in this direction and once the RNA polymerase start and it goes for another 10 nucleotides right. So if it goes for 10 nucleotides then the sigma factor will actually going to be dissociates right. So sigma factor is actually going to dissociate from the temp from the RNA polymerase and then it will actually be available for the next gene and that is how you see in this is event also the bacteria is trying to conserve the energy right. The same sigma factor suppose you have 10 different types of genes and you want to do the transcription right. So sigma factor will go and sit in the gene number 1 right and it will facilitate the process of RNA polymerase to come and start synthesis right.

So it is actually going to start the transcription right. So this is called gene number 1. Now as soon as this is done the sigma factor will come out from here and it will go to the gene number 2 right. Then from here as soon as it is done it would can go to the gene number 3 and so on right and that is how you see that you do not have to synthesize neither the RNA polymerase nor the sigma factor and you can be able to efficiently be able to synthesize the RNA from the DNA modules and that is how you can be able to have the efficient system and the you are at the other hand you are also going to have the conservation of the energy as well. So the first step is the binding of RNA polymerase to the template DNA and the initiation.

The second part is that the binding of RNA polymerase to the template DNA and chain initiation. So the DNA replays should be so this is what anyway we have discussed already. Then you are at this stage it is actually going to form the two different types of complexes. One is called as the open complexes the other is called as the closed complexes. So let us see what happened when it is going to form the open complex.

So open complex which is actually going to form when the sigma factor is going to bind the closed complex is converted into open complex by melting a short region of DNA that is the minus 10 base pair and the RNA polymerase bind at the promoter region and unwind and it covers minus 55 to plus 1. Remember plus 1 is the first nucleotide for the transcription. So total 50 base pairs 55 base pairs and start the initiation here of the one template strand available for the incomplete nucleotide for the base pairing and synthesis of RNA occurs. Minus 10 region of the template is essential for the recognition the promoter regions are double standard in closed complex and single standard in the open complexes. RNA polymerase has two binding sites for the nucleotide one is the initiation site and the other is the elongation site.

Elongation site binds to the first nucleotide within the open promoter complex at the plus 1 site which is usually a purine or the purine A or G. It means the first nucleotide would be either the ATP or the GTP. Elongation site binds with the second incoming nucleotide base pairing at the plus 2 positions. The two nucleotides are joined together and the first base is released from the initiation site and the initiation is complete. So this is what it is actually going to happen.

So first it is actually the sigma factor will go and recognize the promoter and then it is actually going to facilitate the binding of the RNA polymerase and then RNA polymerase when it binds to the promoter region is actually going to unbinds and it is actually going to uncover the 55 base pairs. So it is going to have the plus 1 site. So from minus 55 to plus 1 site it is actually going to unwind the DNA and then it is actually going to start the initiation which means at the first nucleotide it is actually going to add the nucleotide and it prefers that that particular nucleotide should be either A or the G. So that is why the first nucleotide would be either the ATP or the GTP. Then it is going to have the second nucleotide and there will be a bond which is actually going to be formed between the first nucleotide and the second nucleotide.

So it is going to have the bond which is going to be formed and that is how it is actually going to start the synthesis. So this is going to be the initiation step. After the initiation it is actually going to enter into the elongation step. So the chain elongation, chain elongation occurs in the 5 prime to 3 prime direction and RNA synthesis is carried out by the transcription bubble which forms due to the transient separation of double-

stranded DNA into the single-stranded RNA and the transcription takes place at a complex time. So once it is actually going to leave the promoter region it is actually going to enter into the elongation site and that is how the RNA polymerase will be keep moving and it is going to keep synthesizing the RNA.

Now after this it is actually going to reach to a region which is going to be a termination site and that is how it will enter into the termination region. So RNA chain synthesis occur basically at 5 prime to 3 prime ends direction by adding a nucleotide at the 3 prime end and the 3 prime end group of the last nucleotide is combined to the incoming of 5 prime gamma phosphate nucleotide. Alpha and beta phosphate groups are removed and only the gamma phosphate is used in the formation of phosphodiester bond. Likewise other nucleotide ada which are complementary to the template DNA and thus the RNA chain termination strand translocation occurs. In bacterial transcription rate is nearly 40 to 50 nucleotides per second at 37 degree Celsius which is nearly same as the translation in prokaryote which is 50 amino acids per second.

RNA polymerase points to the promoter and create a transcriptional bubble RNA polymerase moves along with the DNA RNA chain grows continuously the length of the transcriptional bubble is approximately 12 to 14 nucleotides. So this is this open area is going to be of 12 to 14 nucleotides and the length of the RNA DNA hybrid is about 8 to 9 base pairs. So within this there will be a region where the RNA and DNA will still have that double standard DNA double strands and it is going to have the RNA DNA hybrid. As the RNA polymerase moves the duplets reforms against the RNA hangs at the free nucleotide chain, free poly nucleotide chain the transcription bubble moves continuously by disrupting the DNA structure. Bioscopic acids are added covalently to the 3 prime on the chain of the RNA beta and drama phosphates are removed from the incoming nucleotide and hadrosil is removed from the 3 prime carbon nucleotide present at the end of the chain.

So that is how it is actually going to occurs into the elongation phase. Now once it reaches to the termination side it is going to end up the into the terminations phase and that is how the termination or the transcription is going to occur. So chain terminations so when the RNA polymerase stop adding a nucleotide at the RNA chain it releases a complete product and the RNA chain get free from the termination sequence. During termination all the hydrogen bond breakdown which holds the DNA RNA hybrid together and when the RNA chain is separated from the DNA again from the duplet. So the nucleotide at which the enzyme stop adding a nucleotide is known as the chain termination side.

So at the chain termination side it is actually going to have the stop the progression of

the RNA polymerase that is the first event. Second there will be a disruption or the breaking of the hydrogen bond between the RNA DNA hybrid and after that the RNA is actually going to fall into the cytoplasm and along with the RNA RNA polymerase is also going to fall and that is how it is actually going to terminate the transcription. And once the termination occurs the DNA which is duplex is going to be reformed. There are two different types of mechanism which are being proposed for the termination one is called as intrinsic termination and the other is called as the rho factor dependent terminations.

So let us discuss about the termination. So intrinsic termination intrinsic termination is being done by the sequence present within the termination side. So these sequences are unique sequences so they will actually going to have the one purpose that to stop the growth of RNA polymerase. If you stop the growth of the RNA polymerase RNA polymerase is a very big enzyme so the DNA RNA duplex what is being formed or the hybrid what is being formed is actually holding the RNA polymerase onto the template. So if there is a growth if there is a you stop the growth of the RNA polymerase then RNA polymerase cannot over remain onto the template and that is how it is actually going to terminate the transcription. So how that this occurs actually is that these intrinsic termination sites are actually going to have the sequences in such a way that it is actually going to form a loop like structure.

You see here you have the AATAGGGCAA like that and on this side also you are going to have the GGA GCCC. So if I show you this it is actually going to form a loop like structure like this it is going to form a loop like structure like this and this one also is actually going to form a loop like structure and because of that and you see that these are the high in GC content and high are AT content which is going to be followed on the other side. So because of that it is actually going to form the stem and the hairpin loop kind of structures and you know that once the hairpin like structure is going to be formed the RNA cannot actually have the possibility because these loops are actually going to have the strong GC content and because of that the RNA polymerase cannot break. So in this mechanism of termination the rho factor is not required and the termination depends on the RNA product. It requires the GC rich hairpin, hairpin structure is followed by 7 U residues.

So RNA DNA hybrid requires the forces for holding the elongation complex together. Thus when the hybrid gets detached it collapses and the elongation complex which causes the termination. In this type of termination the depreciation of the polymerase occurs by destabilizing the attachment of the growing chain to the template. During this process the hairpin structure is formed by the transcription while complementary base pairing.

It includes the palindromic sequences. This stem loop structure include the GC rich region which is followed by the U rich region. So because of that it does not get the enough strength to hold the RNA and on the other hand the RNA polymerase will be going to stop by a strong GC rich region and because of these two events it is actually going to stop the transcriptional activity of the RNA polymerase. So the steps in the transcriptional termination is that the different steps are as follows. Here the two inverted repeat that the GC, GC, GC, GC are present in the DNA template which is transcribed. So nearly the six adenine residues follows the second inverted repeat that is the GC, CC, GC and number three is now inverted repeats are forming a hairpin structure which cause the RNA polymerase.

So this is what I was talking about. You are going to have the hairpin structure and it is been formed because you have a very high GC rich region. This is the high GC followed by the U region and U is actually going to have the low affinity. It is going to have the low affinity because the low U is actually going to have the higher affinity for the template and because of that this cannot withstand the RNA polymerase, cannot hold the RNA polymerase and on the other hand this will not allow the RNA polymerase to cross. So RNA polymerase if it is sitting here it cannot go on this side or it cannot actually break this particular bond. So due to the formation of the stem loop structure and the AU bond get breakdown leads the termination and the RNA molecule get separated.

This is what exactly it is actually going to happen. Now the second method is the row dependent termination. So row dependent termination this type of termination requires a row protein and the row is ATP dependent helicase that disrupt the RNA DNA hybrid. So row is actually having a protein which actually has a very high affinity for the RNA molecules. So it is an essential protein which causes the transcriptional termination. Row protein is a hexamer ATP dependent helicase and it actually sub its subunit contains the RNA binding and ATP hydrolysis domain.

These row proteins firstly bind to the sequence which is present at the upstream of termination site. Each sites are called root RUT sites. These sites are rich in the C residues. C residues row factor followed to the RNA polymerase until it do not catch the RNA polymerase. Row follows the RNA polymerase by its helicase activity which is driven by the ATP hydrolysis.

When the RNA polymerase reaches at the termination site the row protein feeds the structure of the polymerase and when the row factors collapse with the enzyme which causes the termination and the new chain get released. So this is what exactly happens. So row factor is actually binding the RNA polymerase and it is running along with RNA polymerase but when the RNA polymerase reaches at the termination site it speeds gets

slower down and that is how the RNA polymerase actually row actually proteins actually catches the RNA polymerase and that is how they are actually going to dislodge the RNA DNA hybrid and at the end the RNA is going to be RNA DNA hybrid RNA as well as row and RNA polymerase will fall into the heteroplasm and that is how it is actually going to terminate the polymerase and the transcription into the prokaryotes. So this is all about the transcription in prokaryotes. In our subsequent lecture we are going to discuss more about the transcription in eukaryotes.

So what we have discussed we have discussed about the transcriptional unit, we have discussed about what is the coding strand and what is the non-coding strands and we have also discussed about the transcriptional machinery in the prokaryotes where we have the sigma factor and the RNA polymerase and we have done with the previous few slides we have also discussed about the transcription in prokaryotes how the different events are occurring and how the termination is occurring. So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss about the transcription in eukaryotes. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 06
Transcription
Lecture-26 Transcription in Eukaryotic System

Hello everyone, this is Doctor Vishal Trivedi from department of Biosciences and Bioengineering IIT Guwahati. And in the course molecular biology we are discussing about the different aspects related to molecular biology and where we have discussed about the cellular structures. We have discussed about the prokaryotic structure and as well as the eukaryotic structures. Following that we have also discussed about the different types of biomolecules. So, we have discussed about the nucleic acids, we have discussed about the proteins and we have also discussed about the enzymes.

While we were discussing about the nucleic acid, we have discussed about the structure of DNA, we have discussed about how you can be able to isolate the DNA from the eukaryotic and as well as the prokaryotic cell and how you can be able to sequence the DNA and other kinds of related informations. While we were discussing about the RNA, we have discussed about the isolation of RNA. So, we have discussed about the different methods of RNA isolations. Following that we have also discussed about the central dogma of molecular biology where we have discussed about how the multiple processes are related to each other and that is how they are actually been responsible for governing the different events what is happening within the cell.

Now, following that we have also discussed the replications and in the current module we are discussing about the transcriptions. So, let us start discussing about this particular important aspects related to central dogma of molecular biology. Now, as far as central dogma of molecular biology is concerned, it is actually a regulated event where you have the multiple processes linked to each other and they are been responsible for production of the proteins and you know that the production of protein or the enzyme is been responsible for the phenotypic features what are going to be exhibited by the cell. In the previous lecture, we have discussed about the direct transcription in the prokaryotes. So, we have discussed about the transcriptional units, we have discussed about what is the structure of the transcriptional unit in the prokaryotic system and so on.

So, as far as the transcriptional unit is concerned, the transcriptional unit is in the prokaryotes is the polycistronic whereas, in the case of the eukaryotes it is actually the monocistrone. So, and every transcriptional unit is actually having the composition of that it is going to have the promoter, it is going to have the coding sequence and it is

going to have the terminations and all these transcriptional units are been present onto the DNA which is responsible for production of or the synthesis of the ribosomal RNA, tRNA and as well as the messenger RNA. You know that the ribosomal RNA is responsible for the formation of the ribosome and that is how they are actually going to be directly be involved into the protein synthesis. Whereas, the transfer RNA is actually going to be transfer the amino acid, you know that the protein is made up of the different types of amino acids, then messenger RNA is actually been responsible for the decoding the genetic information. Present on DNA right.

So, these three RNA molecules are going to be produced from the DNA by a process which is called as transcription right and in the previous lecture we discuss about the transcription in prokaryotes. So, we have discussed about the different types of events, we have discussed about the initiation, elongation, and terminations. Now, as soon as we will talk about the transcription in eukaryotes, the transcription in eukaryote is going to be more complex because the transcription because the eukaryotic system there is a significant difference between the transcription in prokaryotes versus transcription in eukaryotes. One of the major differences is that in the case of prokaryotes, the RNA polymerase what it is actually going to utilized for what the synthesis of the RNA from the DNA is single or the same type. Whereas, in the case of the eukaryotes, it is going to be different for the different types of cells right.

The second point is that the transcription in the bacteria or in the prokaryotic system is going to occur in the cytoplasm. Whereas, in the case of the eukaryotes, it is actually going to occur inside the nucleus. So, that is why the transcription is going to occur inside the nucleus, then it is actually going to form the transcripts what is going to be formed, it is going to be transported out of the nucleus and that is how it is actually going to be utilized for the translation. So, as soon as the transcription in eukaryote is concerned right. The transcription as I said you know in the case of prokaryotes, it is only utilizes the single type of RNA polymerase, which is been attached to the sigma factor and that is how it is actually going to make an hollow enzyme.

And that hollow enzyme is actually going to be utilized for the transcription of the different types of genes. Whereas, in the case of eukaryotes, you are going to have the different types of RNA polymerase. So, RNA polymerase is the enzyme which is responsible for the transcription. The RNA polymerase of the mitochondria and the chloroplasts are similar like bacteria, because you know that the mitochondria and chloroplasts are the prokaryotic in origin right. All eukaryotic RNA polymerase are multi subunit proteins, which contains three different type there are which there are three different types of RNA polymerase, which is responsible for the transcription in the eukaryotes.

You have the RNA polymerase 1, you have the RNA polymerase 2, and you have RNA polymerase 3. Now, RNA polymerase 1 is been utilized for the synthesis of the ribosomal RNA and it is sensitive it is resistance to the aminidine treatment right. So, it is not going to get affected and it mostly been found inside the nuclei ok. So, within the nucleus you have the region which is called as nuclei right. Then as far as the RNA polymerase 2 is concerned, RNA polymerase 2 is been utilized for the synthesis of the messenger RNA messenger RNA of the different genes right.

It is sensitive or the it is very sensitive for the aminidine treatment or the and it is been found into the nucleoplasm. So, it is not it is going to be present within the nucleus. Then we have the RNA polymerase 3, RNA polymerase 3 is required for the synthesis of the tRNA and it is intermediate between the between the RNA polymer 1 and RNA polymer 2 in terms of the sensitivity to the aminidine. So, it is less sensitive, but it is not that sensitive to the RNA pol 2, but it is less sensitive than. So, it is less sensitive than the RNA pol 2 and RNA pol 1 is anyway not sensitive at all and the RNA pol 3 is also been found within the nucleoplasm.

So, these are the 3 different types of RNA polymerases RNA pol 1, RNA pol 2 and RNA polymerase 3. Now, let us talk about the eukaryotic promoters right. So, each promoter contains some specific situations which get recognized by the transcription factor. Eukaryotic promoter has a longer region than the prokaryotic promoter because it contains all those sequences which are important regarding to initiation. It includes the core promoter elements at which the RNA polymerase get attached and form the initiation complex and also for the efficient transcription it requires an upstream promoter elements which are basically been G plus region and at which the transcription factors are bind.

So, just like the prokaryotic system in the eukaryotic system also you are going to have the initiation you are going to have for after that you are going to have the elongation and after that you are actually going to have the termination and in the initiation the promoter actually a region which is outside the coding region, but within the transcriptional unit is actually going to provide the docking site for the transcription factors. So, these are the some of the important event or important difference between the prokaryotic and the eukaryotic system. Remember that in the in the prokaryotic system the promoter was allowing the binding of the sigma factor whereas, in this case you are actually going to have the binding of the different types of transcription factor and although we are not covering the signal transaction in this particular course, but if you go through with any of the signaling event right for example, if the insulin is binding to the insulin receptor it is actually governing or it is actually activating the the important transcription factors then

these transcription when they go and bind to their respective promoters that is how they are actually going to recruit the RNA pol 2 for the synthesis of that particular gene and the gene product. And that is how they are actually going to have the required enzyme and the required protein into the cytosol and that is how these are actually going to be responsible for the reduction in the glucose level. Same is true for other kinds of transcription other kinds of events for example, if we are getting the infection for example, there will be a COVID infection right then the cells are actually going to have the different types of signaling cascades and that is how it is actually going to activate the different set of transcription factors and these transcription factor will go and bind to the different set of genes the promoters of those genes and that is how they are actually going to have the different types of protein products.

And these protein products are going to be secreted out of the cell and these are mostly been called as cytokines and these cytokines are actually going to fight with the different types of infectious organisms same is true for the antibodies also right when there will be a clonal propagation then there will be a activation of the transcription factors and so on. So, transcription factor actually do play a very crucial role because they decide which promoter they are actually going to bind once the transcription factor goes and binds the promoter then the promoter is committed for the transcription because you have started you have in you have committed this particular promoter for the transcription. So, it is actually you have started the initiation and once the initiation done then the RNA polymerase will have no option but to sit on this and start the transcription right. So, this is been facilitated mostly by the transcription factor that you are actually going to have the transcription factor for the stress responses you have the transcription factors for the lowering the glucose level and so on. So, the set the battery of transcription factor going to decide which promoter or the promoters are going to be labeled for the transcriptional activity and once the promoter is the RNA polymerase is going to sit right mostly the RNA pol 2 is going to sit it is actually going to enter into the elongation phase then you are going to have the multiple cascade of reactions and that is how it is actually going to attach the nucleotides and that is how the elongation will go and then it is actually going to enter into the the termination.

So, termination events are more or less the same as what we have discussed in the prokaryotic system as far as the mechanistic issues are concerned it is going to have the intrinsic termination or the row dependent terminations. So, these are the theme some of the events what we are going to discuss now. So, once the initiation is done then you are actually going to have so during the initiation during the initiation we are going to have a cascade of events. So, that the transcription factor will go and bind and then ultimate aim of the initiation when the initiation is going to happen that the RNA polymerase 2 will come and bind to the promoter and I am always saying RNA pol 2 because RNA pol 2 is

actually been responsible for the synthesis of the messenger RNA and we are mostly having the genome with the different types of genes where the gene product is going to be synthesized. So, mostly it is actually the RNA pol 2 which is going to be transcriptionally very active RNA pol 1 and 3 are only going to be utilized for the synthesis of tRNA or the ribosomal RNA.

So, their activity is going to be not that much compared to the RNA pol 2. So, the initiation so it in transcriptional initiation by the RNA pol 2 eukaryotic messenger RNA transcription required the initiation complex which consists of the general transcription factors and the mediators. So, you are these are the some of the transcription factors which are actually going to be present on to or which are actually going to have the active role into the initiation. So, what we have is the first is this right transcription factor 2D or it is also been called or it is been consist of the tata binding proteins as well as the TBP associated factors or tata binding protein associated factors, but the function of these TF 2 do TF 2D is that it is actually going to recognize to the core promoter or the tata box and it also going to recognize the core promoter which is non-tata box. So, it is going to recognize the tata box and it is going to be differentiate the other transcriptional other other sequences within the promoter.

Then we have the TF 2 transcription factor 2A. So, it is actually going to stabilize the transcription the TBP and that transcription associated the factors binding then we have the TF 2B. So, it is going to help in RNA pol 2 and TF 2F recruitment and also helps in the start site selections. So, these are this is very important actually. So, this is the second factor then you are going to have the third factor then the fourth factor then you have the TF

2E.

So, it helps to RNA pol in the promoter binding then we have the TF 2E. So, TF 2E helps in the TF 2H recruitment modulation of the TF 2H and helicase activity ATPase activity and the kinase activity and then the last is the TF 2H. So, TF 2H is going to help in the promoter melting with the helicase activity and the promoter clearance by the phosphorylation activity. So, you will understand once we are going to discuss how the these transcription factors general transcription factors are going to be play a role in terms of initiating the initiation. So, these initiation is a very you know is a sequential steps and sequential in manner these transcription factor will come and bind to the promoter site and that is how they are actually going to recruit the RNA pol 2 and once the RNA pol 2 is been recruited then the transcription is going to be started.

So, transcription initiation these transcription factors are sequentially going to bind to the TATA box DNA to form a pre initiation complex. At last when the TF 2H get bind its phosphorylate the RNA polymerase to initiate the transcription in the presence of

ATP. So, you have this is a promoter right. So, you have the TATA box and you have the other region in the TATA box right. So, the first transcription factor what is going to bind is the D.

So, TF 2D right it will go and bind then it is actually going to the help or it is actually going to provide the docking site for the TF 2A to bind. So, the first is this is going to bind the second the TF 2A is going to bind and then that once these two are actually going to bind then it is actually going to allow the binding of the TF 2B and then followed by TF 2F E and H as soon as the TF 2E is actually going to provide the docking of the binding of the TF 2H and you know that the TF 2H is actually going to have the helicase activity and also going to have the kinase activity. So, it is actually going to bind or it is going to allow the binding of the RNA polymerase and it is actually going to phosphorylate the RNA pol 2 right you know that the DNA is negatively charged right. So, it is going to be negatively sorry sorry it is going to be negatively charged right. So, when the RNA polymerase is bind RNA polymerase is binding because RNA polymerase is recruiting or binding to this particular region because of the positively charged interaction.

So, there will be a positive charge on to the RNA pol 2 and that is how they are actually interacting with each other. Now, what will happen is once the TF 2H is going to be present what it is actually going to do is it is actually going to convert these positive charges by the phosphorylation. You know that if I have a protein and if I add the you know phosphorylation right. So, imagine that if A is a positive charge right if it is a positively charged protein right that is how it is binding to the DNA. So, that is why it is binding to a DNA.

So, this is not A this is RNA pol 2 and because this is positive and DNA is negatively charged right, but once I will actually going to have the TF 2H right TF 2H is actually going to have the kinase activity it is actually going to phosphorylate. So, what will happen is that it is actually going to generate this right and that is are that is actually it is going to impart the negative charge right. So, if it is imparting the negative charge then it is actually not going to destroy the affinity of the RNA pol 2 to the DNA, but it will actually going to allow the RNA polymerase to move up because it is initially it is binding and it is not moving because that interaction is very strong. Now, it has broken those interactions. So, that it actually can slide over this particular molecule right because the sliding of the RNA polymerase is very important and that is how this is the function of the TF 2H which is actually going to have the helicase activity.

So, it is going to break open the DNA and on the other hand it also going to have the phosphorylation activity. So, that it could be phosphorylate the RNA polymerase. So,

you can imagine that it is going to sit on the DNA and then it is going to slide along the DNA molecule and that is how it is actually going to start synthesizing the RNA molecule. So, it is going to start synthesizing the messenger RNA. So, this is once the initiation at the initiation the RNA all this happens then it will enter into the second phase that is the elongation.

So, elongation is very simple as like what we have discussed in the case of the prokaryotic system where the it is actually going to read the template and then depending upon the Watson-Crick base pairing requirements the A if it is A then it is going to put the T into the sorry U into the RNA into the RNA if it is G then it is going to provide the C remember that there is no T present in the into the RNA structure. So, if it is A then it is going to provide the U if it is G it is going to provide the C. So, this is actually going to be DNA in and this is going to be RNA that we have only discussed also if it is C then it is going to provide the G. So, these are these are going to be depending upon the Watson-Crick base pairing rule and so on. So, if that is the way it is actually going to keep adding and that is how the elongation will be keep happening.

So, from the 5 prime end it is going to start and then it is going to end on to the 3 prime end. Once the elongation is done then it will reach to a stage where it has to stop the synthesis and that is how it is going to enter into the third phase that is called as the trans terminations. So, transcriptional terminations so transcriptional terminations would be different in the case of the RNA pol 1 genes or RNA pol 2 genes and RNA pol 3 genes. So, RNA pol 2 transcription genes may continue to the hundreds or even thousands of nucleotides beyond the end of a coding sequence. Then the cleavage of the RNA strands occur by a complex which appear to be associated with the polymerase.

Cleavage of RNA is coupled with the termination process in occur at the same consensus sequence. The polyadenylation of the mature pol 2 messenger RNA occur at the 3 prime end which results in the poly A tail this process is followed by the cleavage and termination. Both process poly adenylated and termination occur at the both same consensus sequences and both of these processes are interdependent. So, termination is also been governed by the multiple factors and for different types of RNA polymerase you are going to have the different types of terminations. For example, for the RNA pol 1 you are going to have the row dependent terminations whereas, for the RNA pol 2 where you are going to have the most complex terminations that RNA pol 2 termination generally coupled with the RNA processing event in which the 3 prime end of the transcript undergoes the cleavage and poly adenylated.

Whereas, in the case of RNA pol 3 it is actually going to be the row independent terminations. Now, the transcription termination so, you can have the poly A dependent

termination. So, this type of terminations are basically coupled with the RNA maturation process in which the 3' prime end of the nascent RNA undergoes polyadenylation and cleavage and uses the 3' prime end processing reaction as carried out in 2 steps. So, transcription of a poly A followed by the cleavage or nascent and then the upstream product is polyadenylated and downstream product is degraded. Basically the 3' prime starts when the cis acting element in the poly A site of the nascent RNA transcript is recognized by the binding factors.

When these factors bind at 3' prime it form a strong complex result in a high shear force consequently processing down which cause the disruption of RNA pol 2 and DNA RNA. So, this is what exactly going to happen when there will be a transcription in terminations. So, this is going to be a poly A dependent terminations. So, this is all about the transcription in the prokaryotic and as well as the eukaryotic system. So, what we have discussed that from the DNA you are going to have the production of the ribosomal RNA you are going to have the production of tRNA and you are also going to have the synthesis of messenger RNA from in the case of prokaryotic system it is going to be the single RNA polymerase which is going to perform these transcriptional activity with the help of the different types of sigma factors whereas, in the case of the eukaryotic system it is going to be the different types of RNA polymerase which are going to be utilized.

So, you are going to have the RNA pol 1, pol 2 and pol 3 and all of these are actually going to go through with a very complex process of the transcriptional initiation where you are going to have the different types of the transcription factors and these transcription factors are going to be recruited on to the promoter site in a sequential manner and that is how they are actually going to form the pre initiation complex and once the pre initiation complex is going to be formed then it is going to allow the RNA pol 2 to enter into the elongation phase and once the elongation phase is over then it is actually going to enter into the termination phase and the termination is also different for the RNA pol 1 and pol 2 and pol 3 and they are actually going to have the different types of factors which are going to be utilized for the terminations and once the termination is over they are actually going to have the RNA transcripts and these RNA transcripts are further going to be utilized for the protein production or the other kinds of reactions like for example, in the case of RNA polymerase ribosomal RNA or tRNA they are going to be going to be processed for the attachment of the amino acids and other kinds of events whereas, in the case of messenger RNA it is actually going to be utilized for providing the information into the system so, that it is actually going to be utilized for the synthesis of proteins. So, with this brief discussion about the transcription in the prokaryotic and the eukaryotic system I would like to conclude my lecture here in our subsequent lecture we are going to discuss more aspects related to transcription. Thank you. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 06
Transcription
Lecture-27 Post Transcriptional Modifications

Hello everyone, this is Doctor Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati and what we were discussing we were discussing about the different properties of the molecules within the course molecular biology. So far what we have discussed we have discussed about the basic properties of the cells within that we have discussed about the cellular structures. So we have discussed about the prokaryotic and as well as the eukaryotic structures and then we have also discussed about the organelles structures. So we have discussed about the structure of the different organelles. And in the previous module we have also discussed about the different types of biomolecules and what are the different structure and functional properties of these biomolecules. So we have discussed about the DNA, RNA, proteins, enzymes and how the activity of these molecules are being utilized by the cell so that they could not be able to perform the different types of functions.

And in the previous module we have discussed about the discovery of the genomes and how the genome is playing a crucial role in relaying the information from the one generation to the next generation and in addition to that we have also discussed about the central dogma of molecular biology. So when you talk about the central dogma of molecular biology it is a series of reactions which are being required by individual cell or even the organisms to produce the proteins and these events are being tightly controlled and regulated at multiple steps. So in this process we have the three different processes. In the process one you are actually going to have the synthesis of the new DNA from the pre-existing DNA through a process which is called as replications.

These reactions are being catalyzed by the enzyme which is called as DNA dependent DNA polymerase. Then subsequent to that the DNA is also been responsible for production of the or synthesis of the RNA and this process is called as transcription and this process is also been catalyzed by the DNA dependent RNA polymerase and once the RNA is been formed it is actually going to be utilized by the protein synthesis machinery to produce the protein and then this process is also called as translation. And in the current module we are discussing about the transcriptions. So if you recall in the previous two lectures we have discussed about the transcription in the prokaryotes and in the transcription as well as the transcription in the eukaryotes. So when we were discussing about the transcription in prokaryotes we discuss about the different events

like the initiation, elongation and terminations.

Within the initiation we discuss about the how the RNA polymerase is been assembled onto the promoter side and what is the composition of the promoters and so on. So and subsequent to that we have also discussed about the eukaryotic promoter, eukaryotic transcription and within the eukaryotic transcription we discuss about how the pre initiation complex is been formed, how the RNA polymerase is different from the eukaryotic from the prokaryotic RNA polymerase and so on and what are the different events are happening within the eukaryotic transcriptions. So in today's lecture we are going to discuss about the another important topic that is related to the transcription and that is called as post transcriptional modifications. So transcription the generation of the RNA from the DNA is called as transcription and as a result of transcription you are going to have the three different types of RNA you ribosomal RNA you going to have the tRNA and you are going to have the messenger RNA. So formation of RNA from the DNA is known as the transcription and at the result of transcription you are going to have t3 different types of RNA species you going to have a ribosomal RNA you are going to have tRNA and you are going to have messenger RNA.

Now if you recall when we were discussing about the transcription in the prokaryotes the transcription in prokaryotes you are going to have the single RNA species a single RNA polymerase to do this job right. Whereas in the case of eukaryotic transcription you are going to have the three different types of RNA molecules you RNA polymerase molecules you are going to have the RNA pol 1, pol 2 and pol 3 and all these are actually going to have the separate set of the genes which are responsible for the production of ribosomal RNA, tRNA and messenger RNA. Now these messenger RNA and ribosomal RNA are actually going to be presented as a crude molecule ok. So once they are been synthesized they are actually going to be refined or I will say they are actually going to be modified in such a way that they will be more competent in terms of doing the job. So they are actually going to go through with the process of called as post-transcriptional modifications.

So once you have synthesized they are actually going to be present as a crude molecules and then you are going to have the post-transcriptional modification. The purpose of post-transcriptional modification could be very different. One of the major process is that it is actually going to be required for increasing the stability of the molecule. The second is it is actually going to provide the attachment site for many types of molecules. So it is going to be for attachment site and the third is it is actually going to be required for the making the molecules more versatile.

So that it is actually going to be interacting with more number of biomolecules and they

will be going to do many more functions. So first we will going to talk about the post-transcriptional modification within the messenger RNA and then we are going to take up the post-transcriptional modification in the ribosomal RNA and as well as the tRNA. Now messenger RNA so messenger RNA there will be three different types of post-transcriptional modifications what is going to happen in the messenger RNA you are going to have the addition of a cap to the 5 prime end then you are going to have a poly tail at the 3 prime end and then you are going to have the splicing of the introns from the gene. So this these two events are actually been required for increasing the stability of the molecules. So RNA polymerase is sorry so messenger RNA is been synthesized as a crude messenger RNA right and then it is actually going to be modified.

So initially you are going to put the 5 prime cap and you are going to have the coding sequence and then you are going to put the 3 prime poly tail and you are also going to remove the unwanted regions within the gene and that is how it is actually going to be a mature messenger RNA. So this is going to be mature messenger RNA where you are actually going to have so you are going to remove the introns right and you are also going to put the 5 prime cap and you are also going to put the 3 prime poly tail. Now let us first discuss about adding of the 5 prime end cap. So adding of cap to the 5 prime end so 5 prime and cap the as I said you know already that it is actually been required for providing the stability of the molecule. So capping in eukaryotic cell the messenger RNA is in integrably unstable at the end.

So it needs to be modified at the end to protect it against the ribonucleases. Messenger RNA is capped so that it is protected from the ribonucleases as well as it is important in the binding of messenger RNA to the ribosome for the translation. It uses the certain cap binding protein complexes. Capping reaction start soon after the transcription has started right. Remember that the as soon as the transcription start the messenger RNA 5 prime end is actually going to be out right and that is how the capping reaction will start so that the 5 prime end should be get protected from the ribonucleases and it is also going to serve as a docking site for the ribosome assembly and it is going to be assembled on that particular

So that anyway we are going to discuss when we are going to discuss about the translation. So as soon as the 20 to 30 nucleotides are formed the capping occurs. At the 5 prime end the capping process occurs a slightly modified guanine the 7 methyl guanine is attached backward by a 5 prime to 5 prime linkage to the triphosphate of the first transcribed base. Capping reaction includes the condensation of the GTP with the triphosphate at the 5 prime end followed by the methylation of the guanine at 7 side. Further methylation occur at the 2 prime end hydrolysis of the second and third nucleotide adjacent to the cap.

So this is actually the 7-methanine-guanine-cane cap which is actually going to be placed onto the 5 prime end of the messenger RNA and how this is actually going to be synthesized that you are actually going to have the nucleotide. So you are going to have the guanine triphosphate or GTP and from guanine GTP there will be phosphohydrolase enzyme which is actually going to remove the PI and as a result of this it is actually going to remove the gamma phosphate. And then you are going to have this and then this is actually the messenger RNA and then you are going to have the addition of the GTP. So GTP also is going to have the 3 phosphate rings you are going to have the alpha, beta and gamma and then you are going to have the guanine transferase and guanine transferase is actually going to transfer the GTP onto this and as a result and there will be a removal of PPI. So these two are actually going to be removed and these are actually going to be attached onto this and then there will be a release of PPI.

This release of PPI is again going to be in form the 2 PI and that also is going to give you the energy. And then you are going to have this complex and then this complex is going to be get methylated by the guanine 7 methyl transferase enzyme and that is how the there will be addition of methane at the G site right on the guanine site. So there will be addition of methane, methionine methyl group at this site and then there will be another methylation at the 2 prime site and that reaction is going to be catalyzed by the 2 methyl transferase and that is how you are going to have the 7 methyl guanine scene which is actually going to be attached onto the messenger RNA. And this cap is very stable because this cap is not going to be recognized by the RNA polymerase because this cap is not neither because this cap had so many methylations and all those kind of modifications. So they are actually not going to be recognized by the RNA is what is present in the cytosol and then on the other hand this cap is actually going to be specifically be recognized by the ribosome and that is how it is actually going to be responsible for initiating the protein synthesis from this end.

Then the second modification is the adding addition of a poly A tail on the 3 prime end. So on this side you are going to have the poly A tail. So eukaryotic messenger RNA has a series of adenosine residue ranging from the 80 to 250 in the number forming a poly A tail at the 3 prime end of the primary transcripts. This poly A tail has several uses. It can be export mature messenger RNA out of the nucleus site.

It increases the stability of the messenger RNA and number 3 it serves as a recognition of the sequence for the binding of translational factor during the initiation of the translation. The process requires the template independent RNA polymerase called a poly A polymerase. So this is going to be the reactions site. So then the first step what you are going to have is you are going to have the binding of the CPSF. So that CPSF

will go and bind to the primary transcripts.

And then there will be a binding of the additional factors which is called as STSF and CSF and things are actually going to bind the polyadenylation site. And then there will be a binding of the poly A polymerase. Remember that the poly A polymerase is a template independent polymerase. Remember that the DNA polymerase requires a template. RNA polymerase also requires a template.

But in this case the poly A polymerase does not require a template and then the polymer A polymerase is actually going to add the series of A's on to the 3 prime end of the primary transcripts. The length of this poly A tail like for example it can range from the 80 to 250. So depending upon the number of A's it is actually going to say that what will be the age of the messenger RNA. Then we have the third step and the third step is that splicing or the intron removal of the intron. So remember that in a eukaryotic system what you have is you have the genomic region.

The genomic region is actually having the two regions. One is called as the axon and the other is called as the intron. And so you can have the multiple axons and introns and these introns are actually the non-coding region. So these introns are the non-coding region of a non-coding region what is present in the gene. So since these are non-coding region and they will be present within the coding region this non-coding region has to be removed and then only you can be able to attach the axon 1.

So this is the axon 1 and this is axon 2. So you can actually need to attach the axon 1 to the axon 2 by removing the intron through a process which is called as the splicing. So this is been achieved after the RNA is or messenger RNA is been synthesized. So introns are the non-coding nucleotide sequence within a gene that do not code for the protein and do not appear in the final messenger RNA molecule with that is removed by a process which is called as splicing. Protein coding sequence of a gene known as the axon which are interrupted by the intron.

The vast majority of the eukaryotic genes are interrupted by non-coding regions that is intron which needed to be spliced out. However histone protein coding gene is in the vertebrate in a one exception. So in a histone protein coding gene you are actually going to have the nose splicing. The occurrence of the intron varies in the eukaryotic species some each species slacks intron and many genes in the eukaryotes carry the a dozen of them. New bacterial and archal genes also have the introns.

So introns can vary in a length from 50 to 20,000 nucleotides. In higher animals such as human the introns are more than the axon which means in higher animals like mammals

you are going to have the more bigger region of the non-coding region rather than the coding regions. There are four classes of the introns. You can have the group one introns, you can have the group two introns both are the self-splicing intron and does not involve any of the protein enzymes. Then you have the spliceosomal introns they are the not self-splicing introns and then you also have the introns that require the ATP for the splicing.

So depending upon the structures and the other kinds of features the introns could be of four different types. Group one, group two, spliceosomal introns and the intron that do not require the ATP for the splicing. So splicing is a very very important feature, but it is mostly been associated with the eukaryotic genes rather than prokaryotic genes. So as a thumb rule most of the eukaryotic genes are actually having the intron although there are exception that in some of the yeast gene there is no intron present and also the prokaryotic genes do not contain the introns, but there are exception that some of the bacteria and the archaeabacterias are actually having the introns in their genes. So keeping the exception on a side introns are present in the eukaryotic system and introns are absent in the prokaryotic system.

So let us discuss first the splicing by the different mechanisms. So splicing mechanisms first is the splicing mechanism in the group one and group two introns. Splicing mechanism of both the group one and group two remember these are the self splicing introns involves the similar steps of the two trans-terification reaction in which a ribose two prime or three prime hydroxyl group makes a nucleophilic attack onto the phosphorus and the new phosphodiester bond is formed at the expense of the old. So this is the mechanism in the group one and this is the reactions in the group two introns. So what happened is that this is the region what you are supposed to remove.

So this is the intron and intron is always been characterized by a specific pattern of the nucleotides what is present in that particular region. For example if you are having the G and all that. So what will happen is there will be a two event of the nucleophilic attacks by the two prime and three prime hydroxyl group and as a result so what will happen is the first nucleophile attack would be on this side. So this is going to be first nucleophile attack and as a result this bond between the pi prime this bond between the this is going to be broken and then this is actually going to have the nucleophilic attack on this right. This is going to have the nucleophilic attack on this and as a result this particular portion is actually going to be removed and this is actually going to form a bond.

So this is actually exon one and two and this is actually going to be the intron what is present. So in a group one intron the there will be a nucleophilic attack from the this OH right what is present on the interface on the boundary of the first exon and the intron and

then there will be a second nucleophilic attack from this OH on to this phosphate and as a result there will be a bond which is going to be formed and this the in exon intron is actually going to be removed. In the group two where you are actually going to have the lariat formation there will be exactly the same way that you are going to have the two rounds of the transesterification reactions. So and that actually is going to result into a addition of the one and two exon and there will be a removal of lariates.

Then we have the alternate splicing. So the alternate splicing mechanism is a method which is substantially used for many mammalian genes can result in the multiple product that vary structurally and functionally from the same primary transits. Some type alternate splicing is unregulated phenomena which is some in strictly regulated. One of the best example of regulated alternate splicing occurs in the sex determination in Drosophila. In Drosophila three genes are involved in sex determinations sex lethal cell that is the XL transformer gene that is the tra genes and double sex gene that is the ESX. So in the Drosophila what you have or in the alternate splicing what we have is you are actually going to have a gene for example, this gene has four exons followed by the intron.

So you are going to have the exon number one two three and four. Now what will happen is that you are going to have the primary transcript that is going to be formed as messenger RNA. And then you can actually have the multiple combinations you can have the one two three you can have the one three four. So this one is actually going can come along with two and three and that is how you are going to have this. This is the transcript one and if one comes with the three and four then it is going to form this and there are many more other combinations also like for example, one can come with one two and four this is going to be the third combination and you can also have another combination that is the like one followed by three followed by four or one followed by two followed by so one followed by three one followed by four.

So all these combinations could be possible or two three four actually there will be another combination that is like two followed by three followed by four. So these are the some of the alternate splicing where one exon is making a combination with two and three or one exon is making a combination with three four or one exon or two exon is making a combination with three four. So these are the different combinations what could be possible and as a result of this only the mammalian genome has the potential to produce different types of proteins and different types of proteins even from the single gene. Due to the alternate splicing the functional genes are produced in females and non functional genes are produced in the male. So this is a just example that where the one example is in Drosophila where the sex determination is being done by the three genes that is called as sex lethal genes, transformer gene and double sex genes and as a result

of alternate splicing the functional genes are produced in the females whereas, non functional genes are produced in males.

Alternate splicing occurs using two mechanism one when we have two or more poly A or cleavage sites are available in the primary transcripts cleavage occurs at either site resulting in the two products. Such mechanism is followed by the variable domain of the immunoglobulin heavy chain and their diversity is due to the mechanism of alternate splicing. Similarly the alternate splicing with such mechanism results in the product of two different hormones calcium regulating hormone in the red thyroid and the calcitonin gene related peptide in the red brain. Other mechanism involves more than three prime site or five prime site hence splicing occurs by taking either of those three prime splice site resulting in the different products.

Then we have the another messenger RNA splicing. Such mechanism is followed by a variable domain of the immunoglobulin heavy chain and other variant is due to the mechanism of alternate splicing. So the alternate splicing is a very robust phenomena what is happening within the eukaryotic system and it actually allows the production of different types of variations within a biological system and as a result you are actually going to have the different types of protein what is going to be produced even from the single gene. Now let us move on to the next RNA that is the ribosomal RNA. So processing of ribosomal RNA eukaryotes have 80S ribosomes whereas the prokaryotes have the 70S ribosomes. Ribosomal RNAs are transcribed as a long precursor sequence which is then modified at a specific basis and cleavage to give the mature product.

In both bacteria and eukaryotic ribosomal RNA processing involve two basic steps of cleavage and the base modifications. So ribosomal RNA processing in bacteria. So ribosomal RNA precursor in bacteria is a 30S RNA ribosomal RNA which is modified and cleaved to give 23S ribosomal RNA, 16S ribosomal RNA, 5S ribosomal RNA and some tRNA segment in between are also there sometime. 30S pre ribosomal transcript consists of the 16S ribosomal sequence followed by the spacer which may have tRNA sequence in some cases and there is a 23S ribosomal RNA sequence followed by the 5S ribosomal sequence near 3 prime end. At times there is one more tRNA sequence after the 5S ribosomal RNA sequence at the 3 prime end.

There are several different genes for ribosomal RNA in E. coli they are essentially similar in sequence of ribosomal segment, but differ with number and sequence of tRNA segments. Maturation process involves the methylation of the 3 prime 30S ribosomal precursor at a specific site occurring at 2 prime hydroxyl group of bases. Some bases such as uridine is modified to pseudo-uridine or dihydro-uridine. Further cleavage

process is carried out using the enzyme RNase 3, RNase P and RNase E at the site 1, 2 and 3 respectively.

Intermediate products are formed mainly as 70S ribosomal RNA, tRNA, 25S and 5S. These are acted on by certain nuclei to give the final product of 16S tRNA, 23S, 5S respectively. So this is actually going to be the primary transcript where you are going to have the 16S ribosomal RNA, tRNA gene, then you are going to have 23S ribosomal RNA and the 5S ribosomal. And after this you are going to have the mature RNA where you are going to have the 16S ribosomal RNA, you are going to have the tRNA, you are going to have 23S ribosomal RNA and you are going to have the 5S ribosomal RNA. And you know that all of these are actually going to come together to give you a 70S ribosomal RNA.

Then ribosomal RNA processing in the eukaryotes. So in the eukaryotes nucleolus is a site of processing the ribosomal RNA. A 45S ribosomal precursor is formed by the RNA pol I and processed in the 90S pre ribosomal nuclear complex to give the 18S, 28S and 5.8S ribosomal RNA. There is a tight coupling of RNA processing with the ribosomal assembly 5S ribosomal RNA is transcribed by the RNA pol III from a separate gene. Precursor RNA undergoes methylation at more than 100 bases from the 14,000 nucleotide at 2 prime and hydroxyl groups.

Furthermore, there is a modification of bases such as uridine to serodilidine etc. followed by a series of cleavage reactions. Cleavage and modifications are guided by the small nuclear roller RNA in yeast the entire processing involves a pre RNA, 170 non ribosomal protein, 70SNO RNA and 78 ribosomal protein. SNO RNA are supposed to be the remnant of the nucleosomes. So this is going to be the ribosomal RNA pre ribosomal RNA which actually contains the gene for 18S, 5.

8S and 28S and then this is going to be cleaved and from separately you are going to have the 5S ribosomal RNA. And these all going to be combined and it will actually going to give you a mature ribosomal RNA that is the 80S ribosomal RNA. Now let us talk about the processing of the tRNA. So we do you do not have to worry about the structure of the tRNA because that anyway we are going to discuss when we are going to discuss in the into the because the tRNA has a major role in the protein production. So that we are going to discuss when we are going to discuss about the translation.

So in both the eukaryotic and prokaryotic tRNA processing occurs it is transcribed as a long precursor sometime as a single primary transcript carry more than tRNA one tRNA segment which is separated by the cleavage. The processing of tRNA involved cutting off the extra sequences by the endonucleases such as RNase P at the 5 prime end and

RNase D at the 3 prime end. RNase P is a ribozyme which is a RNA exhibiting catalytic activity. After removal of the sequences from the 3 prime end the CCS sequence is added by the enzyme tRNA nucleotide transferase. This enzyme binds to the sequence at its active site and phosphodiester bond is formed with the 3 prime end.

Furthermore there is a base modifications occurring simultaneously such as methylation, deaminations or reduction. In case of the pseudo uridine the uracil is removed and reattached to the sugar through the 5 end. So when you will see that there are different regions in that within the tRNA you are going to have the different types of CCA and you are going to have all those kind of things you are going to have anticodon chain and all that. So that time you will be able to understand why the tRNA is actually going to be modified for all these modifications and that is how you are actually going to have the mature product and these mature product are actually going to participate into the protein synthesis machinery. So what we have discussed so far we have discussed about the replications we have discussed about the transcriptions and in the current module we have discussed about the prokaryotic transcription we have discussed about the eukaryotic transcription and in the current lecture we have discussed about the post transcriptional modifications.

So within the post transcriptional modification we discuss about the how the messenger RNA is actually going to be modified so messenger RNA is going to be capped at both the ends it is actually going to have the 5 prime cap and it is going to be protected from the 3 prime end by having a poly A tail apart from that the messenger RNA is also going to be modified by removing the introns and this process is known as the splicing. So there are going to be 4 different types of splicing introns you are going to have group 1 introns you are going to have group 2 introns you are going to have the spliceosomal introns and you also require the ATP. So these are the 4 different types of splicing different types of introns what are present in the eukaryotic system. In general the introns are not present in the prokaryotic system and they are only present in the eukaryotic system but there are exceptions both in the prokaryotic system and as well as in the eukaryotic system. For example there are yeast genes where there is no introns present and there are prokaryotic bacteria or the RK bacteria where the introns are present.

So these are some exceptions but in general the introns are absent in the prokaryotic system and introns are present in the eukaryotic system. Now introns are actually playing a very crucial role because it is allowing the alternate splicing and alternate splicing is a robust mechanism and a tool through which the eukaryotic organisms are actually been able to produce the multiple types of different types of proteins by the single transcripts.

So within the transcripts the multiple types of genes can come together right for example if within the one transcript you have the 4 genes 1 2 3 4. So either the 1 2 3 can come together or 1 3 4 can come together and that is how you are going to have a protein which is made up of 1 2 3 or 1 3 4. So from the single transcript you can have the different types of proteins.

So and at the end we have also discussed about the both transcriptional modification in the ribosomal RNA and as well as tRNA. You will be able to understand more when we are going to discuss about the ribosomal RNA structure of the ribosomal RNA and the structure of the tRNA in the subsequent modules and why there will be a modification required in these structures. So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 06
Transcription
Lecture-28 Gene Control Mechanism (Part 1)

Hello everyone, this is Dr. Vishal Tewedi from department of bios assessment by engineering IIT Guwahati. And in this particular course, we are discussing about the different aspects of the molecular biology. So, for what we have discussed, we have discussed about the basic structures of the cell. So, we have discussed about the prokaryotic structure and eukaryotic structures. Following to that, we have also discussed about the different types of biomolecules, we discussed about the different types of cellular activities.

And in the previous lecture, we have discussed about the central dogma of molecular biology and we have also discussed about the different components, which are be a part of the central dogma of molecular biology. So, we have discussed about the replication in detail about in the previous module. And we have discussed about the replication in the prokaryotes and the replication in eukaryotes. And then we also discuss about how the replication is helping the cell to recover from the different types of cellular damages or DNA damages actually.

In the current module, we are discussing about another important aspect related to the central dogma of molecular biology and that is the transcription. So, if you recall in the previous few lectures, we have discussed about the transcription in eukaryote prokaryotes followed by the transcription in eukaryotes. And in the previous lecture, we have discussed about the how the post-transcriptional modification is happening in the different types of RNA species, what is going to be produced from the DNA and then how these modifications are enabling these RNA species to work optimally for the protein production. Now, if you see the central dogma of molecular biology, what you see is that the ultimate goal of the central dogma of molecular biology is that it is going to produce a protein or I will say protein or the enzyme actually. And these proteins or the enzymes are actually going to participate into the different types of the metabolic reactions.

For example, when we are discussing about the carbohydrate metabolism in the eukaryotic cells, we discussed that the carbohydrates are actually as soon as the glucose enter into the cell, it get phosphorylated by the hexokinase followed by the phosphorylation by the different types of followed by the glucose phosphate and other

kinds of molecules and ultimately it is going to channelize that particular glucose molecule into the glycolysis. So, if the glucose if you see about glucose metabolism glucose will enter and then it is actually going to be processed by the different types of enzyme to produce the pyruvic acid. And these cascades of the reactions are going to be called as the glycolysis which occurs within the cytosol and pyruvic acid then will enter into the Krebs cycle and ultimately it is actually going to produce the different amount of energy and it is also going to produce the different types of intermediates which are actually going to be utilized for the different types of synthesis. So, Krebs cycle is going to provide the raw material for the synthesis of the different types of amino acids. Krebs cycle is going to provide the raw material to synthesize the hemine which is very important component for the blood synthesis or in general I will say hemoglobin synthesis and then Krebs cycle is also important for the synthesis of the fat and as well as the nucleic acid.

So, all these components are important because they are actually going to be required for running the different types of the life related activities. Now imagine that when the glucose level is down right or there is no supply of glucose for example, the glucose from the glucose from where your glucose will come the glucose will come from the food and as well as the glucose will enter into the cell it will go through with these kind of reactions, but if there will be access of fat then these reactions will go in this direction right and this will go in this direction and there will be less utilization of glucose and if there will be less utilization of glucose the enzymes which are being present within the glycolysis are going to be down regulated because there is a only when mechanism there are several mechanism what we have discussed when we were talking about the glycolysis or the Krebs cycle how the Krebs cycle and the glycolysis is being regulated, but apart from the feedback mechanism or allosteric regulations there is a also another level of regulation is that you are going to have the regulation at the level of the protein synthesis which means you are actually going to make the availability of these enzymes or the protein at lower level or the higher level and depending upon the amount of these proteins or the enzyme that particular type of activity is actually either will go up or will go down actually. So, these kind of events are more relevant when you are talking about the bacterial system because the bacterial system is going to have the polycystronic you know the transcriptional unit right compared to that eukaryotic system has the monocystronic transcriptional unit. So, when you have polycystronic transcriptional unit this means you are going to have the different types of enzymes being produced simultaneously from the single transcript and in that case all the protein synthesis of all these enzymes are actually going to be under tight regulation or tight control. So, that you will be able to have the complete control over the different types of events such as glycolysis, Krebs cycle and all other kinds of things right.

Although the Krebs cycle is not present in the bacterial system or the prokaryotic system but for the sake of examples there could be say many more other kinds of pathways for example you are going to have the fatty acid synthesis pathway and you are going to have the other kinds of pathways like amino acid biosynthesis pathway and so on. So, all these pathways require a very tight control and one of the mechanism through which the bacteria is actually bringing the control is by up regulating and down regulating the amount of protein what it actually going to synthesize. And this is all been achieved by putting these enzymes or putting these genes or putting the transcriptional unit under a complete control mechanism right and all these are been a part of the operon which means a system which actually is going to operate or going to regulate the transcription of these particular transcription followed by the synthesis of these proteins. So, a typical bacterial cell contains several thousand genes some genes carry out the universal task and are constantly active these are called as the housekeeping genes. For example, the housekeeping genes include those that facilitate the synthesis of the protein and the ribosomal RNA.

The majority of the genes however only become active when their byproducts are needed such genes should not be expressed constitutively because the energy could be could be used for more productive task. So, what is mean by the constitutively means that it is actually going to be expressed throughout the life cycle it is not going to be induced right. So, you can actually have the two different two modes of the expression one is called as the constitutive the other one is called as the induced. So, induced means in case you have some action some kind of thing right for example, when you stand in the sun you are going to have the sweating that. So, that sweating reaction are induced reaction because they are being induced by the sunlight whereas, you are going to have the constitutive reactions constitutive reactions means you are going to have the irrespective of whether you are in sun or light or whatever you are going to have that for example, the running of our heartbeat right.

So, our heartbeat is going to be constitutive reactions. So, such gene expression is controlled. So, the induced gene expression is controlled. So, their products are only produced when they are required in accordance with the need of the cell the phase at which the gene expression can be controlled are numerous. So, in prokaryotes the most common step at which the regulation of gene expression occur is the transcriptional initiation right.

You remember that we have said that in the prokaryotes the transcription and the translation go together so, there is no regulation at the translational level in the in the case of prokaryote as soon as the RNA is being produced it is been present in the cytosol right. So, and it is actually going to be taken up by the translational machinery and there

will be a synthesis of the protein. And what you can actually control is the initiation part right as if you do not allow the RNA polymerase to go and bind to the promoter and sit and start the transcription then it is actually going to be controlled. So, it is energetically the most efficient step to regulate the gene expression. The transcriptional regulation occur in step after initiation specifically during elongation and termination.

Prokaryotic transcriptional regulation is accomplished by the gene regulating protein that bind with the regulatory sequence near the transcription site of the transcriptional unit. Gene regulatory protein the product of gene regulatory protein are of 2 types they can be activators and they can be repressor. So, gene regulatory proteins or the gene regulatory components are actually going to be the regulatory units which are going to regulate the efficiency of the RNA polymerase to go and sit on to the transcriptional initiation site and that is how they are actually going to control the transcription. These gene regulatory proteins or the protein gene products could be activator. So, that actually going to activate.

So, they can actually be able to facilitate the RNA polymerase to go and bind to the initiation site or they could be repressor. So, that if they are actually going to block. So, activator means they are actually going to activate repressor which means they are actually going to block or they are going to inhibit the process. In the absence of both activator and repressor transcription carried out by RNA polymerase is called as the basal level of transcription. The binding of a repressor decreases the transcription less than the basal level.

So, repressor is inhibitor right it is going to inhibit whereas, the binding of an activator increases the transcription which is the above the base level. If both the both repressor and activator represent and functional the action of the repressor typically overtake over that the transcription. The basic concept of how gene regulation occur at the transcriptional initiation in bacteria are provided by the classical model called the operon model. This is formulated by the Jacobson Monnet in the 1961. So, how these repressor and the activators are regulating the gene expression profiling and how they are actually regulating the transcriptional activities within the bacteria is being provided by a classical models and these models are called as the operon model and these models are hypothesized and formulated by the Jacobson Monnet in the year of 1961.

So, the question comes what is operon? So, operon is a set of genes. So, operon is a genetic regulatory system mostly seen in the prokaryotes and the bacteriophage in which a group of structural gene are transcribed together under the control of a single promoter, which means the operon technically the operons are over sleeve in present in the polycistronic transcriptional unit and they are going to be present in the prokaryotes and

the bacteriophage where you are going to have the group of structural genes or I will say the genes which is going to be transcribed for the different types of enzymes or the structural genes and they will be under the control of a single promoter. This means this single promoter is actually going to have the control over the synthesis of these structural genes. So, you are going to in a typical operon you are going to have the operon promoter next to the promoter are going to have the operator and the objects to the operator you are going to have this is, this is going to be the coding region. So, this is going to be the coding region and this coding region is going to be responsible for.

So, this is the coding region, and then you are going to have the poly A site. So, the coding region is going to be responsible for for the synthesis of the A protein, B protein and the C protein. So generally operands are very common in prokaryotes and the bacteriophage but it is also found in some eukaryotes. The main difference is that the expression of prokaryotic operand leads to the polycystronic messenger RNA whereas the eukaryotic operands leads to the monocytic messenger RNAs. In this particular lecture we are mostly been focused on to the operand what is being present in the prokaryotic system.

We have not discussed about the operands in the present in the eukaryotic system. So the idea is that you should we should be able to tell you the concept of the transcriptional activations and transcriptional regulations and how the things are been done similar kind of things are also been done in the eukaryotic system. So prokaryotes are the single cell organisms lacking a true nucleus and the membrane bound organelles. It adopt the operand system as a mechanism to efficiently regulate the gene expression in response to the changing environment conditions. Environment condition means the requirements of the different types of metabolites, availability of glucose, availability of oxygen and so on.

So the bacteria has since bacteria is a single cell organisms it actually gets affected very oftenly and it is actually has to respond to these changes. The operand system is a genetic regulatory system found in the prokaryotic organism that allowed the multiple gene with the related function to be controlled as a single unit. The system offers different advantage for prokaryotes like it is energy efficient because you are supposed to only synthesize the operator or the repressor and that actually is good enough to control and regulate the active transcriptional activity of several genes. So operand system allows them to coordinate the expression of multiple gene involved in a common pathway and transcribe together a single messenger RNA saving the energy and resources by producing the necessary protein only when required. Then it is actually going to have a rapid response to the environmental changes.

So the operand system enable them to adapt to the changing conditions quickly. If condition changed the expression of the relevant gene can be turned on or off rapidly. It is simple and compact right. So prokaryotes can use a single regulatory region to control the expression of multiple genes. This is particularly advantage in a small genome where saving space is very crucial.

So if you require a multiple genes as a regulatory genes and so on you are actually going to increase the size of the genome and that the bacterial system cannot afford because bacterial system has to conserve the energy, conserve the space and all those kind of things and that is why it is important that it should actually operate and control the multiple genes with the help of the operands. Then it has the coordinated regulations. It allows the bacteria to go with the coordinated regulation because it can regulate these three genes four genes five genes which are actually been present in the single pathway. For example if you are talking about the glycolysis you are going to regulate the hexokinase you are going to regulate the pyruvate kinase and so on. So since the all these genes are present in a single operand probably it is easy for the bacteria to manage all these and on the other hand it is also going to save the energy.

So there will be a coordinated regulation. So this coordinated regulation ensure that the product of genes these are produced in the appropriate stoichiometric ratio which means if you are processing the single glucose molecule you require one glucose molecule of hexokinase you require the one glucose molecule of aldolase and so on. So you can actually be able to produce these proteins and enzymes in a right proportion so that you should not waste the energy by producing some one or other in excess amount and on the other hand you should not have the lower production of any of these proteins. Then you have the resource allocations when a particular nutrient is available the genes required for the utilization are switched on. Once the nutrient becomes scarce the operation operand can be turned off preventing the wasteful production of the unnecessarily proteins.

Then we have the adaptation to the niche environment so the operand system enable the prokaryotes to adopt in a specialized niche by turning fine tuning the expression of genes that are specifically relevant to those conditions. So this is also very very important that you are actually having a very very fine and regulated balance and that is how you can be able to have the fine tuning of the expression of genes that are relevant that particular for that particular environmental conditions. Then you have the evolutionary advantage organisms with ability to regulate gene expression rapidly and efficiently in response to the environmental changes were more likely to survive and reproduce. So operand as I said you know is been proposed by the Jacobin monit and it is been control mechanisms through which the prokaryotic system is controlling the different types of gene and gene

expressions within the bacteria. So in 1965 the Nobel Prize in physiology and medicine was awarded jointly to the Jacobs Monet for discovery concerning the operon and the viral synthesis.

So this is the all scientists who have got the Nobel Prize in medicine in the 1965 for their concept of the operon. So before we get into the detail of the operons and how we are going to take up some of the examples of the operon it is important to understand what will be the structure of the general structure of an operon. So this is just a general structure of the operon you are going in the general structure what you are going to have. So this is the transcriptional unit right this is the transcriptional unit where you are going to have the promoters you are going to have the operators you are going to have the structural genes for example in this case this is a structural gene for A B and C and apart from that you are also going to have the regulatory genes. So regulatory genes are actually going to be a part of is going to produce the regulatory proteins and going to produce the regulatory proteins and these regulatory proteins could be activators right or it could be repressor and condition of the activator and repressor will go and bind to the operator right and so this region what you see here is actually a part of operon.

So regulatory gene is not going to be a part of operon and regulatory proteins are going to either activate or either going to facilitate the binding of RNA polymerase or it is actually going to have the other way around right. So regulatory gene encodes for a protein called regulatory protein which either act as a repressor or activator which control the operon but it is not a part of operon because it has its own promoter right. So regulatory genes are not a part of operon this is the part of operon where you are going to have the promoters operators and genes for the structural genes. Now let us talk about the regulation of an operon. So there are two types of transcriptional regulation which is going to be possible in the operon one is called as the first is called as the negative control and the second is called as the positive control.

So then the negative control in which the regulatory protein is act as a repressor which binding to the DNA and inhibiting the transcriptional of the protein right. So and then you have a positive control in which the regulatory protein is acting as an activator which is stimulate for the transcription. So this is the regulatory gene from which you are going to have the regulatory proteins and regulatory proteins could either be a repressor which means once it goes and bind to the operator it will not allow the RNA polymerase to go and bind to the promoter and that is how it is actually not going to allow the transcription of so there will be no transcription of the structural genes. This is going to be called as the negative control whereas in the case of positive control you are going to have the regulatory proteins which are actually go and bind to the operators and that is how they are actually going to facilitate the efficient binding of the RNA polymerase to

the promoters and that is how it is actually going to have the more production of the particular structural genes and these are called as the positive regulations. It means when you have this you are going to have the more production when you have this and you have going to have the lower production this means it is going to be negative regulation this is going to be a positive regulations.

Within this you are going to have the two different types of conditions either it going to be inducible or it going to be repressible. Then you going to have so operon can also be either inducible or repressible. So inducible operons are those in which the transcription is normally been off which is not going to take place and it needs inducer to induce the transcription which means it is going to be transcom it is going to be turn on. Repressible operons are those in which the transcription is normally been on which means there will be a basal level of transcription which is going to take place. Sometimes it may happen to repress the transcription or turn it off.

So you are going to have the positive control, negative control and then you can also have the inducible operon or the repressible operon. So this is what I have summarize here. So you are going to have the negative control. You are going to have positive control. And within the negative control or the positive control you can be having the inducible operon or the repressible operon.

So, in the negative control the product of the regulatory gene inhibits the transcription in a positive control the product of regulatory gene is going to activate the transcription. Whereas in the inducible operon your initial condition or I will say the basal level of transcription is going to be off which means you are not going to have the transcription of that particular gene. But once this inducible inducer is present then you are going to have the operon which is going to work. So, it is going to have the turn on the transcription whereas, in the case of repressible initial condition the basal level there will be a transcription, but when the repressible is present there it is actually going to turn off the transcription. So, let us first discuss about the negative inducible operon.

So, within the negative you can have the inducible you can actually having the so you are going to have the inducible or you are going to have the repressible. Even within the positive control also you can have the inducible or the repressible. So, there are several different conditions in which all these has to be understood. So, let us first take the first example that is the negative congruent inducible. So, in a negative so negative inducible operon in a negative inducible operon the regulatory gene encode a repressor which readily binds to the operator as operator side overlap with the promoter side.

So, that the binding of the repressor physically block the binding of RNA polymerase on

the promoter and prevent the transcription. So, for the initiation of transcription something is needed to prevent the binding of the repressor at the operation side and represent the operators site of binding that is the inducer. This type of system is said to be inducible since transcription is usually off and must be turned on. So, in this negative inducible operon what will happen is that you are going to from the regulator you are going to have the repressor. So, this is the repressor molecule which will go and fit and sit on to the operator.

So, since the repressor is sitting on the so if the no inducer is present this repressor will be keep binding to the operator molecule and it will not allow the transcription of the structural gene because repressor will bind to the operator and it will inhibit the transcription. So, there will be a transcription offs. And when the inducer is present what will happen is that the inducer is suppose the insulin for example, or I will say glucose. If the glucose is present what will happen is that the glucose will go and bind to this repressor. And in that case it is actually going to make the active repressor to inactive repressor and then the active repressor would not be able to bind the operator and as a result it is actually going to allow the transcription of these particular structural genes.

So, this is an example of or the mechanism in which the negative inducible operon is going to operate. We are going to take the few examples and then you will be able to understand this more nicely. And then we have the second condition. Second condition is that you are going to have the negative repressible operon. So the regulator gene in this type of operon synthesizes and the inactive operon that cannot bind to the operator.

So, RNA polymerase readily bind to the promoter without any inhibition and transcribe the structural genes. To turn the transcription something must be needed to make the repressor active. A small molecule called a corepressor binds to the repressor and make it capable of binding to the operator. So, in the absence of inducer the regulatory genes are producing the repressor but these repressors are inactive which means they will not be able to bind the operator and that side there will be a transcription.

So, transcription is on right. So transcription is on sorry transcription is on right under the basal level because you RNA polymerase will go and bind to the promoter there is no inhibition because the repressor what you are producing is inactive and that is how there will be a production of on right. So, there will be a transcription on. When the inducer is present the inducer will go and bind to the repressor and that is how it is going to convert the inactive repressor into active repressor and active repressor will go and bind to the promoter and that is how it is actually going to inhibit the transcription and that is how it is actually going to have the turn off it is going to turn off the transcription. So,

this is the another example or another way in which the operand can be can be regulated.

So, this is called as negative repressible operand. Then the third condition is the positive inducible response right. Remember that in the positive it is going to be transcriptionally off and then it is going to be on when the inducer is present. So in a positive inducer inducible operand transcription is usually turn off because the regulatory proteins that is the activator is produced in an inactive form right. Remember that when you are talking about the positive regulation it is going to be inducer it is going to be activator rather than repressor. So, the negative control it is going to be done by the repressor whereas, here it is going to be activator.

So, whatever we have discussed in the case of negative operands it is going to be exactly the reverse. So, in this case the activator is produced, but that activator is in the inactive form which means it cannot activate the transcription right. So, transcription take place when an inducer has become attached to the regulatory protein rendering to the regulatory side. So, when the inducer is not present the regulatory region is producing an activator, but this is a inactive regulator which means it requires some kind of modification. So that it will go and bind to the operator and that is how it actually can enhance the production or enhance the transcription.

So, there will be a transcriptional off right because the this activator is not competent enough or efficient enough to induce the transcription. So, inactive activators cannot activate the transcription and that is why there will be no transcription. But once the you add the inducers these inducer will go and bind to the activators and once the activator binds to the inducers they will actually going to have there could be a structural changes within the activator and that is how they will be actually go and bind to the operator right. And the active operator is stimulate the transcription right and that is how you are going to have the transcription with of the structural genes. Then we have the positive repressible operand this is exactly the opposite of the negative inducible operand.

So a positive operand can also be repressible the regulatory protein is producing an activator and that will bind to the DNA meaning the transcription usually take place and has to be repressed. Transcription is inhibited when a substance become attached to the activator and render it unable to bind the DNA. So transcription is no longer stimulated. So this is exactly the in the case of no inducer you are going to have the active activators and active activator will go and bind the operator and that is how there will be enhanced production of the these particular genes. But when the inducer will be in added inducer will go and bind to the transcription making them inactive transcription activator and inactive activator will actually going to turn off the transcription.

So these are the four different conditions in which the operand can be regulated by the repressor or the repressor proteins and it could be inducible or it could be repressible. So these are the just the summary of what we have discussed so far. So you going to have the repressible operand or you are going to have the inducible operand. Repressible operand generally in normally keep that GLS synthesis on but can be turned off by the repressors whereas in the inducible operand generally the genes are tough with that means the transcription is off but can be turned on by the inducer. Repressible operand are mostly been present in the anabolism reactions or anabolic reactions whereas the inducible operands are always been present in the catabolic reactions.

In repressible operands you are going to have the inactive form whereas in the inducible operand you are going to have the active forms and the examples of the repressible operand is the tryptophan operand whereas inducible operand which is the example of the lac operand. So we are going to take up these examples so that it will be easy for you to understand what is mean by the inducible operand what is mean by the repressible operand and so on. So we have taken first example that is the lac operand and then we are going to take the tryptophan operand. So this is the example where you are going to have the synthesis whereas here you are going to have the breakdown of the substance. So this is going to be related to the catabolic reactions and this is actually going to be related to the anabolic reactions.

So let us first start with the lac operands. So lac operand or the lactose operand which is in short is called as lac operand. So the lac operand of the E. coli contains the gene which are involved into the lactose metabolism. It is expressed only when the lactose is present and the glucose is absent.

This is very important. If you have a glucose in the lac operand will no longer be active. Lactose can be broken down by the E. coli but it is not their preferred energy source. They would much instead use glucose if it is available.

Lactose can be broken down more slowly and with less energy than glucose. However if lactose is the only sugar present E. coli will immediately use it as a fuel. The lac operand contains the 3 structural genes lacZ, lacY and lacA. So these are the 3 genes lacZ, lacY and lacA and all these 3 genes have their own individual roles. So lacZ is called as the beta galactosidase, lacY is called as beta galactosidase permease and lacA is actually been called as beta galactosidase trans-acetylase.

These genes are transcribed as single messenger RNA under the control of the single promoter that is the this promoter. The lac operand is typically been present as a shut off or repressed in a normal condition but can be activated in the presence of the inducer

which is called as lactose or allolactose. Plus the lac operon referred to be an inducible operon. So allolactose is a structural analog of the lactose. So these are the structural genes, regulatory genes and the regulatory DNA sequences and the regulatory gene which are present in the lac operon.

So you are going to have the 3 different types of structural genes lacZ which codes for the enzyme beta galactosidase, the cleaves lactose into the glucose and galactose this enzyme also converts the lactose into allolactose. Then we have the lacY which encodes for the beta galactosidase permease which transports the lactose into the cell. So basically the lacY is so this is if this is the cell the lacY is actually going to bring the lactose into the cell. Okay and then lactose is going to be converted into glucose and galactose right by the enzyme which is called as lac by the gene product of lacZ right. So it is going to be called as beta galactosidase and lacA codes for the enzyme which is called as beta galactosidase trans-tylase.

It is not essential for the lactose metabolism but appears to play a role in the detoxification of the compound by transferring and the acetyl group. Then we have the regulatory DNA sequences remember that in a transcriptional unit you have the promoters you are going to have the coding region and then you are going to have the 3 prime poly A tail right. Apart from this you are also going to have the because we are talking about the operons. So in this case you are going to have the operators. So you are going to have the promoters followed by operators followed by structural genes followed by the poly A tail right.

So the regulatory DNA sequences you are going to have the promoters. The promoter is the binding site for the RNA polymerase which initiates the transcription of the structural gene. Lac promoter is a weak promoter. So remember that when we were talking about the transcription we discuss about the weak and a strong promoters. So there are compositions which actually becomes the which makes the promoter as a weak promoter or the strong promoters right. Because a strong promoter allows the efficient transcription of the DNA right and it allows the efficient very efficiently the RNA polymerase to go and bind the and complete the transcription whereas in the case of weak promoters the melting of the DNA or the other kinds of activities is very difficult and that is how it is actually going to have the lower efficiency and lower production of the RNAs.

Apart from that you are going to have the operators. So the operator is a negative regulatory site bound with the lac repressor protein. The operator overlaps with the promoters. Then we have the cap binding site. The cap binding site is a positive regulatory site that is bound by the catabolite activator protein or the cap. When the cap

is bound to this site it first promotes the transcription by helping the RNA polymerase bind to the promoter.

Apart from that you are also going to have the regulatory genes which is called as lac I. So the regulatory gene lac I transcribed and produced the lac repressor protein and inhibited the lac repressor transcription in order to accomplish this it binds to the promoter partially overlapping the operator. When bound the lac repressor gets into the way of RNA polymerase and prevents the operon transcription but when the lac repressor binds with the lactose it becomes the repressed. There are multiple conditions. So you are going to have the in the absence of so just to make it comparable what we have discussed.

So in absence of inducer so in absence of inducer so remember that lac operon is lactose is going to be an inducer. So in absence of inducer so in absence of inducer or in presence of inducer. So when the lactose is not available which means the inducer is not available the lac repressor strongly bind with the operator and it stop the RNA polymerase from the initiation of transcription. However, lac repressor loses its capability to bind DNA when lactose is present it leaves the operator and float away making it possible for RNA polymerase to transcribe the gene. So if the lactose is not present then the what will happen is that the repressor will not be able to bind the operator.

And so if the lactose is not available the repressor will go and bind to the operator and since the repressor is binding to the operator it will not allow the RNA polymerase to go further to start the transcription. So there will be no transcription of the structural gene of lacZ Y and Z and when the allolactose is absent the repressor bind with the operator so the transcription cannot initiated by the RNA polymerase without any prevention. Now when the lactose or the allolactose is available so what will happen is these inducer will go and bind to the repressor. So when they will go and bind to the repressor allolactose or the lactose it will they will no longer be able to bind the operator and as a result what will happen is that the RNA polymerase will move and it will actually going to do the synthesis of the structural genes and that is how they are actually going to produce the beta lactosidase and other kinds of enzyme from these genes. So when the allolactose bind with the lac repressor, the repressor cannot bind an operator so transcription initiation by RNA polymerase without any prevention.

So some of the apart from the lactose or the allolactose some of the allolactose analogs can also be used for the for the lac promoters or the lac operons one of the very popular lactose analog is the IPTG or isopropyl beta D1 thioglycosidase it is actually a inducer for the protein production and that we are anyway going to discuss when we are going to

discuss about the molecular cloning. Then we have the phenyl beta D galactose which is called as phenyl galactose and then you also have thiomethyl galactose or the TMG. So all these are some of the lactose analog so either the lactose allolactose or these analogs can be able to modulate the activity of the repressor and that is how they can actually be able to have the effect on the lac operon. Then the lac promoter is a weak promoter it does not bind RNA polymerase more efficiently on its own it would not be able to accomplish much more without the help of the catabolite activator protein. High transcriptional levels are facilitated by the CAP binding to a stretch of DNA right before the lac operon.

So this is the CAP binding region and this is the CAP proteins and CAP proteins are actually going to bind by the cyclic AMP. So E.coli produces a cyclic AMP as a hunger signal in the low glucose conditions by attaching to the CAP cyclic AMP modifies the structure of the CAP enabling it to bind the DNA and stimulate the transcription. CAP is inactive without cyclic AMP only when the glucose levels are low the CAP levels are very high does CAP actually activate. So in the condition of low glucose when the glucose levels are low you are going to have the large quantity of the ADP and ADP is going to be converted into AMP and this AMP is actually going to be converted into cyclic AMP.

And the cyclic AMP will go and bind to the CAP region right. So it is actually going to bind the CAP proteins and once they bind to the CAP protein they are actually going to you know they are actually going to block the activity. So when the CAP attached to the CAP cyclic AMP attached to the CAP and activates it allowing it to bind the DNA CAP helps RNA polymerase to bind to the promoter resulting in a high level of transcription right. So in a state of starvation you are going to have a very high amount of ADP and that ADP is getting converted into AMP and then the AMP is getting converted into cyclic AMP and in the case of low glucose this cyclic AMP will go and bind to the CAP and as a result it actually going to activate and allowing it to bind the DNA and CAP is actually going to help the promoter to bind to the promoter and that is how they are actually going to have the high level of transcription of these genes. When there will be a high glucose so in the case of high glucose there will be no production of cyclic AMP and that is how there will be no binding of cyclic AMP to the CAP and as a result the CAP will not going to help the RNA polymerase to bind to the promoter and that is how there will be a low level of transcription.

So there will be high amount of lac operon transcription are only possible without glucose. This method ensure that the bacteria only activate the lac operon and begin using the lactose after exhausting their primary energy source that is the glucose. So if the glucose is present it is actually going to block or it is going to inhibit the lac operon

activity simply because it is actually going to does not allow the production of cyclic AMP and cyclic AMP is going to bind the cap region or the cap proteins and that is how they are actually going to facilitate the binding of RNA polymerase to the promoter. So if we summarize all these conditions what is the conditions you are going to have the glucose absent lactose absent. So there will be no transcription. So there are going to be four conditions in four conditions you are going to have in first condition you are going to have the glucose absent lactose absent in those conditions there will be no transcription of the RNA the lac operon or there will be no lac operon activity because the lac operon inducer is also absent and the glucose is also absent.

So they still there although there will be a production of cyclic AMP because the glucose is absent but since the lactose is also absent the repressor protein will actually going to repressor repress the activity or repress the production of the production of RNA from the RNA from the RNA polymerase. Now the there will be a second condition. So second condition would be that the glucose is absent but the lactose is present in that condition there will be a high transcription because in the absence of glucose we discussed already that from the ADP it is actually going to form the AMP and from the AMP it is actually going to form the cyclic AMP and the cyclic AMP will go into the cap and it will go and bind to the cap region of the DNA and that is how it is actually going to facilitate the promoter. And since the lactose is present lactose will also going to bind the repressor and that is how it is actually going to remove the repression and as a result it is actually going to have the allow the RNA polymerase to go for the transcription and that side there will be a high transcription level.

Now the third condition is that you are going to have the glucose present and lactose absent. So if the glucose is present it is actually not going to allow the production of cyclic AMP and there will be a low level of cyclic AMP so that the cap will not going cap proteins will not be able to bind to the cap region and that is how the RNA polymerase will not be efficiently be able to bind to the promoter and that is how and on the other hand since the lactose is absent the repressor will actually going to bind and it is going to you know allow the operator right. So it is going to allow the repressor to bind to the operator and that side there will be no transcription. In the fourth condition fourth condition both the biomolecules both the molecules are present which means the glucose is also present and lactose is also present in that case there will be a low level of transcription because since the lactose is so if the glucose is present there will be no cyclic AMP so it is not going to efficiently allow the binding of the cap proteins to the cap region of the DNA and that is how there will be a very low level of transcriptional activity from the RNA polymerase because the lactose is present it will actually going to bind the repressor and that is how it is actually going to destroy the inhibition of the operator and that is how it is going to allow the RNA polymerase to go. But this level of

RNA transcription would be less compared to the transcription activity what we have just observed in the condition number two. So if we summarize all these activities summary would be that the lac operon contains the gene those are involved into the lactose metabolisms.

Lac operon is a negative inducible operon the genes are expressed only when the lactose is present and the glucose is absent remember this is very important because the glucose is the primary metabolite and or primary metabolite preferred by the cell whereas the lactose is a secondary metabolite and is not been preferred by the cell. So it will not go to utilize the lactose until the glucose is absent and we have already discussed that when the glucose is absent it is going to have the synthesis of the cyclic AMP and the cyclic AMP will go and bind to the cap proteins and that is how the cap protein is actually going to help into the binding of RNA polymerase to the promoter site and that is how it is actually going to have the higher production of the protein synthesis. The operon is turned off in a normal conditions the operon is turned on and turned off depending upon the glucose and lactose level the catabolic activator protein and the lac operon. The lac operon blocks the transcription of the operon by binding with the operator in the absence in the presence of lactose it stops acting as a repressor. So catabolite activator protein act as an enhancer activate the transcription of the operon but only when the glucose levels are low.

So these are the four conditions what we have already discussed if we have the glucose we have the no lactose then there will be no transcription because the glucose is present and it is going to inhibit the activity of the cyclic production of cyclic AMP. Then you can have the both glucose and lactose present then it will be having the low level of transcription then you third condition where the glucose is also absent and lactose is also absent then there will be no transcription and then the fourth condition is that the glucose is absent but the lactose is present. So in that case there will be a production of cyclic AMP and the cyclic AMP will allow the binding of the cap proteins to the cap region of the RNA polymerase to the promoter and that is why there will be a high level of transcription. And because the lactose is also present it will actually going to block the repressor and lac I and that is how it is actually going to have the high level of transcription.

So this is the summary of the lac operon. So with this I would like to conclude my lecture here in a subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 06
Transcription
Lecture-29 Gene Control Mechanism (Part 2)

Hello everyone, this is Dr. Vishal Tewedi from department of biosciences and bioengineering, IIT Guwahati. And in this particular course, we are discussing about the different aspects of the molecular biology. Now, let us move on to the next operon and the next operon is the tryptophan operon. So, tryptophan operon is a part of the anabolic operon, compared to that the tryptophan operon is a catabolic operon. So, the tryptophan operon

 finds in the E.

coli, it is a group of genes that encodes enzyme for the synthesis. So, remember that the first operon but we have discussed is for the breakdown of the lactose it is going to break down the lactose into glucose and galactose and that is how it is going to be drive the energy from the lactose molecule whereas, here you are actually going to consume the energy. So, this is actually an anabolic pathway, it is a anabolic or negatively controlled operon it always remain on in the normal conditions and off when the tryptophan level is high. So, this is exactly reverse what we have just discussed for the lacto operon.

Tryptophan does not need to be synthesized by the E. coli bacteria when it is present in the environment. Hence, the transcription of a gene in the tryptophan operon is turned off on the other side, when the availability of tryptophan is low the operon becomes on and the genes are transcribed by synthetic enzyme for the tryptophan synthesis. TIP repressor does not always attach with DNA instead it binds and inhibit transcription only in the presence of tryptophan. Tryptophan binds to the repressor molecule and alter their structure which switches an inactive repressor into a active state thus the tryptophan act as

 a corepressor.

Remember that this is very important tryptophan act as a corepressor because it enhances the repression activity of the repressor. So, it become converts the inactive repressor into a active repressor. One unique feature of the tryptophan repressor is the attenuation. So, like regulation by the TIP repressor attenuation is a mechanism for reducing the expression of a tryptophan when the level of tryptophans are high. However, rather than blocking the initiation or transcription attenuation prevent the completion

 of the transcription.

So, this is a very unique feature of the tryptophan operon. Now, the structure of the

tryptophan operon. So, there are 5 structural genes TRIP E, D, C, B and A that code for the enzyme involved in the conversion of the carboxylic acid to the tryptophan. Remember that we have already discussed about the tryptophan biosynthesis when we are talking about the amino acid metabolisms. So, TRIP E actually codes for the enzyme anthonyl synthase 1, TRIP D actually codes for the enzyme anthonyl synthase 2, TRIP C it encodes for the enzyme 5-phosphoryl anthonylate isomerase and the endole 3-bristol phosphate synthase.

Then TRIP B is encodes the enzyme tryptophan synthase B subunits and TRIP A is actually going to encode for the enzyme tryptophan synthase A subunit. The controlling site in the tryptophan lies next to the TRIP E and consists of promoter and overlapping operator and a leader region or the TRIP L. So, in these are the regulatory region. So, this is the operon and this is the regulatory region, you are going to have the promoter, you are going to have the operators and you are going to have the leader region which is called a TRIP L and then you are going to have structural genes like TRIP E, TRIP D, TRIP C, B and A and they are actually going to give you a polycystronic messenger RNA. So, it is also contain a repressor regulatory gene called TRIP R.

When the tryptophan is present the TRIP R protein binds to the operator blocking the transcriptional tryptophan operon by inhibiting the RNA polymerase binding. So, this is the repressor regulatory gene TRIP R. So, when it is going to be produced it is going to bind the tryptophan and that is how it is actually going to make the active repressor and that is how it is actually going to allow the binding of the repressor to the operator region and that is how it is actually going to block the transcription of the these genes structural genes. Reactions catalyzed by the enzyme synthesis from the tryptophan operon. So, this is anyway we have discussed in detail the biosynthesis.

Ultimately from the indole you are going to have the synthesis of the tryptophan where are you going to have the activity of enthnynyl synthesis, enthalase, transferase, PRA isomerase, IGP synthase and so on. So, all these genes are actually going to be of part of the tryptophan operons. This we are not going to discuss in detail because we have already discussed these things when we were talking about the tryptophan biosynthesis. Now, tryptophan operon regulations. So, in the absence of tryptophan which means then when the bacteria actually require the synthesis of the tryptophan.

So, there will be low tryptophan into the environment. When there is a little tryptophan or absence of tryptophan in the cell in this condition the tryp repressor is inactive because there is no tryptophan available to bind with the repressor and activate it by the conformational change. So, inactive repressor cannot bind to the DNA that is the operator or block the transcription which allows the tryptophan to be transcribed by the

RNA polymerase. So, once there will be an absence of tryptophan which means there is a low tryptophan present then the repressor is not going to be active because it has to bind the tryptophan molecule to become active repressor and that is how it will not be able to bind to the operator and as a result the RNA polymerase will go and bind and it is actually going to do the transcription and it will actually going to produce the polycystonic messenger RNA. So, there will be high level of gene production or the transcripts from the operon.

Now in the presence of tryptophan when there will be high tryptophan the things are going to be reversed because if the tryptophan is present it will go and bind to the repressor proteins and as a result it is actually going to form the active repressor and if the active repressor is going to be present it will go and bind to the operator and that is how it will not allow the RNA polymerase to bind to the promoter and to complete the transcription and that is how there will be a low transcription in the case of the if the tryptophan is present. Then we have the tryptophan operon transcriptional attenuation. So, it is possible to obtain more strict regulation in E. coli by repressing the transcription initiation alone but translation mediated transcriptional attenuation offer the additional regulation. The attenuation site in the tryptophan operon is situated after the surfactational start site.

More transcription stop here when the tryptophan levels are high when the tryptophan levels are low or sparse thus transcription continue to produce the functional protein. So, in this case you are actually going to have the attenuation site. So, this is actually going to be attenuation site what is present. So, in the case in the high tryptophan level it is actually going to be transcribed and it is actually going to produce a messenger RNA which is for this particular protein. So, when the tryptophan concentration is low the entire operon including the ligur sequence is transcribed into a messenger RNA.

When the tryptophan concentration is high only the 140 nucleotide which is only the part of the sequence that precede the attenuator are transcribed into messenger RNA and the structural genes are not been transcribed. So, it is actually when you have the low level of tryptophan the transcription will start from here and it will go all the way up to the end. So, it is actually going to have the full length messenger RNA where you are going to have the leader sequences and as well as the RNA sequences of the structural genes. Whereas when you have a very high level of tryptophan the transcription will start from here but it will only going to end up here and you are only going to have the leader sequence of 140 nucleotide and that is how you are actually going to stop this in the transcription of these particular genes. And this is a very unique phenomena what is only happening in the tryptophan operon which is called as transcriptional attenuation.

So, what is the mechanism? So, the operation leader sequence has a 14 codon open reading frame codes for a leader peptide of 14 amino acid which with 2 tryptophan codons. The mechanism of translation mediated attenuation depend on the fact that the translation in bacteria is coupled with the transcription. So, ribosome becomes translating the 5 prime end of the messenger RNA when it is still being synthesized. Thus the translation rate can affect the structure of a growing RNA chain which determine whether the transcription can continue or not. The function of the leader sequence is to fine tune the expression of tryptophan operon based on the availability of tryptophan inside the cell.

The 2 tryptophan codes for the leader sequence lies within the region 1 and the transcriptional translational stop codon lies between the region 1 and 2. The leader sequence contain the 4 region that is the region 1 to 4 and that can form the various base paired stem loops or the hairpin like the secondary structure. So, regions are like you have the region 1, region 2, region 3 and region 4 and region 3 is complementary to both the region 1 and region 4. And as a result what will happen is that it is actually going to form a hairpin like structure. So, it is actually behaving exactly the same as we have discussed about the intrinsic transcriptional stop site.

So, it is actually going to form a loop kind of structure and as a result it is actually going to stop the growth of the RNA polymerase. If the region 3 and 4 base pair with each other they form a loop like structure called attenuator and function as a transcriptional terminator. If pairing occurs between the region 3 and 2 then no such attenuation forms and the transcriptional continues. So, this is the exactly the site what we have just discussed that ribosome binds to the tryptophan polycystin messenger RNA that is being translated when the tryptophan levels are high and start the leader sequence translation. The translation stop codon is present between the region 1 and 2 and the 2 tryptophan codes for the leader sequence are within the region 1.

The ribosome follow the messenger RNA closely during translation and creates the leader peptide. This peptide the moving ribosome complete the translation of the leader peptide and pause at the stop codon blocking chain blocking region 2. At this point the ribosome prevents the region 2 from interacting with the sequence 3. So, the base pair with the region 4 to form a 3 4 stem loop which serve as a transcriptional terminator and as a result tryptophan prevents the tryptophan operon from continuing to be transcribed. So, this is what exactly happened when you have the high level of transcription tryptophan.

So, there will be a 2 tryptophan coding region what if been present and they are actually going to allow the formation of a loop like structure and this loop will actually going to

stop the progression of the RNA polymerase and as a result it is actually going to stop the RNA polymerase and stop the transcription. So, this is exactly what we have discussed. So, if this tryptophan is in short supply then the ribosome will pause at the 2 tryptophan codon contained within the sequence 1. This leaves the sequence 2 free to the base pair with sequence 3 to form the 2 3 structure also called as anti terminator. So, the 3 4 structure cannot form and transcription continues to the end of the tryptophan operon.

So, when the tryptophan levels are low there will be a base pairing of 2 and 3 and this 2 and 3 are called as the anti terminator because it will not be able to find a strong loop structure and that is how the transcription will continue and that is how it is actually going to have the synthesis of the messenger RNA for the tryptophan synthesis. Now let us move on to the third operon and the third operon is called as the arabinose operon or the ara operon. So, the 5 carbon sugar L-arabinose must be breakdown by a operon known as L-arabinose operon also known as the ara or the ara-B bad operon in the E.coli. The 3 structural genes ara-B, ara-A and ara-D are found in the L-arabinose operon code for the 3 metabolic enzyme needed for the breakdown of L-arabinose.

These genes generate the enzyme called arabinose or the ribulokinase, ara-A which is called an isomerase and ara-D which is called an epimerase which catalyze the conversion of the L-arabinose into the D-xylose 5 phosphate and intermediate in the pentose phosphate pathway. So L-arabinose is also a part of the catabolic pathway and it is going to follow exactly the same mechanism what we have just discussed about the lac operon. So, a single transcript and messenger RNA is produced from the transcription of all structural gene in the L-arabinose operon. The catabolite activator protein or the cap cyclic AMP complex which is produced by the regulatory gene ara-C regulate the expression of L-arabinose operon as a whole. The proteins that codes ara-C control the expression of arabinose by acting as both activator when the arabinose is present and a repressor when the arabinose is absent.

ara-C is sensitive to the level of arabinose at high ara-C level the ara-C protein not only regulate the expression of arabinose but also control its own expression. So, these are the metabolic pathway of L-arabinose by the action of 3 enzymes. So, when you have the L-arabinose it is going to act by the L-arabinose isomerase and that is going to convert the L-arabinose into L-Ribulose and then L-Ribulose is going to act by the L-Ribulose kinase and that is how it is going to form the L-arabinose 5 phosphate and L-arabinose 5 phosphate is going to be isomerized by the L-arabinose epimerase and that is how it is going to produce the D-xylose 5 phosphate. So, what is the structure of the arabinose operon? So, you are going to have the regulatory genes like ara-C you are going to have the promoter which is called as para-C and then you are going to have the different types

of structural genes like ara-B, ara-A and ara-D apart from that you are going to have the some of the regulatory proteins and all that. So, this is the region of the L-arabinose operon where the ara-B is going to be produced by for a ribulose kinase, ara-A is going to form the isomerase and then ara-D is going to form the epimerase.

L-arabinose operon consists of three structural gene and the regulatory region with the region with the operator region called ara-O, ara-O1 and O2 and the initiation region that is called as ara-I1 and I2. So, these are the region right the structural genes are ara-B, ara-A and ara-D and this is also there is also a cap binding site where the cyclic cap and cyclic MP complex bind to and facilitate the catabolic repression and result in the positive regulation of ara-B when the cells lack the glucose. The regulatory gene ara-C is located upstream of the L-arabinose operon and encodes the ara-B responsive regulatory protein ara-C both ara-C and ara-B have a specific promoter where RNA polymerase bind and initiate the transcription ara-B and ara-C are transcribed in opposite direction from the ara-B promoters and ara-C promoter respectively. Now ara-Bnose operon regulations in addition to being under the control of cap cyclic AMP activator the ara-Bnose system is also positively or negatively regulated by the binding of ara-C proteins. So, ara-C perform as a homodimer and interact with the operator and initial range region of initiation range of the ara-Bnose operon to control the transcription of ara-B. ara-B a DNA binding domain and a dimerization domain make up such ara-C monomers 2 domain.

So this is the DNA binding domain so this is ara-Bnose binding site and this is the structure of ara-C monomer and there will be a dimerization. So, you are going to have the ara-Bnose binding site and you are going to have the DNA binding site. The binding of ara-Bnose is carried out by the dimerization domain upon binding to ara-Bnose ara-C undergo and the question will shift and adopt 2 different confirmations the binding of the allosteric inducer ara-Bnose is also only factor that affect the confirmation. When the concentration of ara-C rises to high ara-C can potentially adverse auto regulate its own expression by attacking dimer ara-C to the operator region and ara-C production is in between. So, now you have the negative regulation of the arabid.

So, cell do not require the arabid product to metabolize ara-Bnose when it is not present. So, in the absence of ara-Bnose you are going to have the negative regulation. So dimeric ara-C therefore function as a repressor 1 monomer binds to the ara-B genes operator while the other monomer bind to the remote DNA half site called Arab A. DNA loop is created as a result the Arab ad promoter cannot be bound by the RNA polymerase while in this operation orientation as a result structurally Arabid transcription is blocked. So, this is what exactly going to happen you are going to have the operators which are actually going to dimerize and that is how you are going to have the binding of the

operators onto the Arab 2 and Arab 2 onto the DNA and that is why it is actually going to form a loop like structure and in this loop like structure the RNA polymerase will not be able to bind.

Then you have the positive regulation of Arabid. So, both in the presence of Arabonose and the lack of glucose the Arabonode operator is activated for expression. So, when you have the low glucose you are going to have the ADP followed by AMP followed by production of cyclic AMP and that is how you are going to have the cyclic AMP cap proteins complex formation and that complex is going to bind the cap region of the DNA and on the other hand when the Arabonose is present Arabonose will go and bind to the operators and that is how it will not allow the interaction of the operators to form the loop like structure like this and that is how there will be a transcription of the structural genes from the operons. So, ARAC and CAP cooperate and act as a activator when Arabonose is present. So, when glucose is absent a high level of cap protein cyclic AMP complex bind to the cap region side binding of cyclic AMP is responsible for opening up the DNA loop between the era 1 and era 2 O2 increasing the binding affinity of RAC protein for era I2 and thereby promoting the RNA polymerase to bind to the Arabate promoter to switch on the expression of the Arabate required for Arabonose metabolism.

So, this is all what we have discussed in relation to the regulation of the transcriptional regulation through the operons. Now, what we have discussed we have discussed that the operons are being functional mostly into the prokaryotic structure, but the operons are also present into the eukaryotic structures and in a typical operon what you have is you have a promoter then you followed by the operators and followed by the structural genes and these operons or can be positively been regulated or the negatively been regulated or they can be inducible or the repressible. So, in this context we have discussed about the 2 operons from the catabolic reactions and 1 operon from the anabolic reactions. So, in the lac operon it is a catabolic operon where you are going to have the mostly the operon is going to be present as the negatively regulated operon. So, there will be no production of the proteins and there will be no production of enzymes, but when there will be an absence of glucose and there will be a presence of lactose then the bacteria will actually going to have the transcription of these genes and because the repressor is going to be bind by the lactose and that is how it is actually going to relieve the inhibition and that is how there will be a transcription of the gene and that gene is actually going to act on to the lactose molecule and that is how the lactose is going to be get converted into glucose and galactose and that is how that glucose will be utilized by the glycolysis to produce energy.

Apart from that we have also discussed about the tryptophan operon and we have also discussed about the arabdron subchloron. So, this is all about the discussion about the

operons and how the operons are actually regulating the transcriptional activity within the prokaryotic system. So, with this I would like to conclude my lecture here in a subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 07
Translation
Lecture-30 Translation in Prokaryotic System

Hello everyone this is Dr. Vishal Tewedi from department of bioscience and bioengineering IIT, Guwahati and in the course molecular biology we are going to discuss about the different aspects of the cell biology we have discussed about the cell divisions different types of organelles what are present in the eukaryotic cell and then we also discuss about the different types of biomolecules. So, we discuss about the nucleic acid proteins enzymes and then subsequent to that we have discussed we are in last couple of modules we are discussing about the central dogma of molecular biology. So, within the central dogma of molecular biology so far what we have discussed we have discuss about the replications transcription. So, we have discuss about the replication different steps of replications in the prokaryotic system and the eukaryotic system and we have understood that how the machinery is actually unwinding the DNA how the DNA is being replicated and it is making the two copies of the DNA and what are the different types of events are happening within the prokaryotic system or the eukaryotic system. Similarly, we have also discuss about how the DNA is giving rise to the RNA through a process known as transcription and the transcription is also very different in the case of prokaryotic versus eukaryotic and in the prokaryotic system the transcription is been is much simpler compared to the eukaryotic system.

So, we have also discuss about the different types of differences between the prokaryotic system as well as the eukaryotic system. Now in today's lecture and in this particular module we are going to discuss about the translation. So, translation is the third step which is going to be where the RNA what is been produced within the transcription is going to be utilized for the production of the protein right and you know that the major reason why we are studying the central dogma of molecular biology because the central dogma of molecular biology is the basis of the life on the any kind of organisms or in the on the earth itself. Because all these events like the replication transcription or translations are you know tightly regulated and these events are tightly linked to each other.

So, if you want to you know make that two copies of a particular cell you are supposed to do a replication, you are supposed to do a transcription, you are supposed to do a you know the translation. Because until you were not going to do that replication the one DNA copy will not going to be you know make the two copies right. So, how you are

going to divide the genomic content between the two daughter cells. Then at the end at the same time when you are actually going to produce the next cell you are supposed to provide him the sufficient amount of the you know or different types of organelles you are supposed to provide him the different amount of other kinds of nutrients and other things. And for that you require the first the production of RNA through a process of transcription and then you are also going to have the another process of translation.

So, that you are actually going to synthesize the different amounts of proteins and enzymes and all other kinds of biomolecules. So, that you can be able to run the metabolism you can be able to govern the different types of events and so on. And you know that all these processes and all these events are very crucial for maintaining the life on the earth actually. So, that is why we are studying the central dogma of molecular biology. Now let us come back to the our topic.

So, in today's topic we are going to discuss about the translation. And when we say translation, translation means that you are actually going to synthesize proteins. And if you recall when we were discussing about the protein we said that the protein is made up of the amino acids. And if you want to synthesize the protein the amino acid has to be joined together. So, I hope that you remember what are the reactions are happening.

So, for example, you can imagine that if this is the amino acid 1 and if this is the amino acid 2 they are actually going to combine with each other with to through a peptide bond. So, this is a peptide bond what is present here. This is a peptide bond and then there will be a dehydration reactions. So, this is actually a standard way of synthesizing a protein with the help of the amino acids. But in the biological system you actually require the help from the different types of species or different types of system.

Because this under the in vitro system you can actually be able to choose the amino acid 1, you can be able to choose the amino acid 2 and that is how you are actually going to get a dipeptide. Where the first amino acid would be in the first place and second amino acid will be in the third place. You can even think about synthesizing the similar kind of protein with the help of the multiple amino acids. So, you can have amino acid 1, 2, 3, 4, 5, 6 I like that. But in biological system when you want to synthesize a protein the machinery supposed to you know select the amino acid 1, supposed to select the amino acid number 2 and so on.

So, how that works? So, that purpose the DNA is synthesizing the different types of RNA species. So, it is synthesizing the ribosomal RNA, it is synthesizing the transfer RNA and it is synthesizing the messenger RNA. All of these RNA species are being synthesized from the DNA with the process known as transcription that anyway we have

discussed in our previous class right or in the previous module. So, the why there is a reason why we are actually synthesizing the all these three different types of RNA species right. Because of the simple reason that we do not know the sequence of the amino acid in which we are supposed to attach right.

How in which amino acid will come first, then second, then third, then fourth like that. So, that sequence we do not know, but that sequence information is available on to the DNA right. Because DNA is actually the information system right. It is actually the hard disk or the brain of the particular cell. So, it actually knows which cell which protein which if you are synthesizing a protein X which amino acid will come first which amino acid will.

So, that information it is giving to the messenger RNA and that is why it is called as the messenger RNA right. So, it actually going to convey or it is actually going to take away the message from the DNA. And that message is nothing, but the different types of nucleotides right. So, it is actually going to have different types of nucleotides sorry. So, it is actually going to have the different types of nucleotides like A T A C G just like that ok.

So, this is actually going to be derived from the DNA. This process is also going to be done through a process of transcription. Now, once you have these then it is you are providing the information into the system, but who is actually going to read these information. So, that you can be able to know which amino acid. So, for example, this amino acid is actually going to read in pairs right and we are going to discuss in detail why it is been pair discussed in pair and this reading is actually going to be done by the tRNA.

So, tRNA is actually a amino acid where you have the two parts you are actually going to have we can imagine that it is actually having a two reading blocks on one or two other right. So, it is going to have the lower end which is actually going to sense the codes what are present on to the messenger RNA and on the top side it is actually going to have the amino acid what is present as the as like amino acid is attached covalently to this. So, this is actually a tRNA right. So, tRNA is actually going to have two parts and the lower end you are going that is actually going to help to read the codes what are mean what is been derived from the DNA right and the upper side it is actually going to have the corresponding amino acid right. For example, if this is for methionine then you are actually going to have tRNA which is going to be for methionine ok.

Same is true then this whole thing is and then they are actually going to do this condensation reactions right. So, that condensation reaction is actually going to be done

in the ribosomal RNA. So, ribosomal RNA is actually going to prepare more protein complexes and that is actually going to be called as ribosomes and ribosome is having a small subunit and it is going to have the large subunit and within these large and small subunit the tRNA is actually going to provide these amino acids and that is how these amino acids are actually going to combine with each other and through a condensation reactions and from the ribosomes you are actually going to have the synthesis of the proteins right. So, you can see that how complicated it is how systematic it is right and where the messenger RNA is actually you know conveying the or taking the message from the DNA tRNA is reading that message from the messenger RNA and as well as it is bringing the appropriate amino acids and then it is taking those amino acid into the ribosome which is actually a multimeric protein complexes involving the different types of ribosomal RNA and then it is actually going to do a condensation reaction just like this right. So, amino acid 1 and amino acid 2 is combining with each other to make a dipeptide.

Similarly this reaction can continue for several rounds and that is how you are going to have the synthesis of even a bigger protein like 300 amino acid, 400 amino acid, 500 amino acid protein right. So, this is just a brief overview of the importance of the different RNA species into the whole process of translations. Now if you want to understand the whole process of translation we are supposed to understand the many aspects related to messenger RNA, many aspects related to tRNA and many aspects related to ribosomes. So, what are the things we supposed to understand first? So, first thing what we have to understand is the structure of translational machinery which means we should understand the structure of the ribosomal RNA, we should understand the structure of the ribosomes right. So, then only you will be understand how the you know tRNA is getting into that ribosome how the tRNA is you know how the ribosome is catalyzing the condensation reactions and then you also should understand the structure of tRNA.

So, that it we will understand how it is binding to the amino acids, how it is recognizing the messenger at the course on the messenger RNA and then we also should understand the structure of the messenger RNA and very briefly we discuss about the messenger RNA structure. So, that you can be able to understand much quicker in the when we were going to discuss about this and then we also going to discuss about the genetic code remember that when we were talking about the messenger RNA the nucleotide sequence what is present on to the messenger RNA which is been derived from the DNA that is a you know genetic code which you are going to read and that is going to be read by the tRNA. Then we are going to discuss about the mechanism of translation. So, mechanism of translation involves the activation of amino acids, initiation, elongation and termination. This is a mechanism of translation could be you know could be for the

prokaryotic system or could be for the eukaryotic system.

So, remember that we are discussing these kind of mechanisms for both the systems. So, prokaryotic as well as the eukaryotic system. Although the basic steps remains the same that you are going to have the activation of amino acids. So, that amino acid going to attach to the tRNA and then you also going to have the different steps like initiation, elongation and termination all these are actually going to happen on to the ribosomes. And then so basic steps remains the same whether it is a prokaryotic system or eukaryotic system, but depending upon the complexity of the system depending upon the different types of requirements and different types of adaptations the prokaryotic system is adopting the different strategies to you know to perform these events or eukaryotic system is adopting the different.

Remember that the eukaryotic system is much more diversified much more you know wealthy in terms of the energy and in terms of the other kinds of resources and that is why it is actually doing the things in a very very different way compared to the eukaryotic system. So, let us start the first thing the first thing is the structure of the translational machinery. So, the structure of the translational machinery where we are first going to talk about the messenger RNA. So, messenger RNA this all we have discussed before also. So, this is going to be a revision for you.

So, messenger RNA has a 3 prime 5 prime end has a 5 prime end 5 prime UTR ribosomal binding site coding sequences and 3 prime UTR. In eukaryotes there are additional structures such as 5 prime one in cap and poly A tail actually. So, technically the messenger RNA will have the 5 prime end 5 prime UTR these are going to provide the binding of the different types of factors then it is also going to have the ribosomal binding site then you are going to have the coding sequence this is the coding sequence which is going to be responsible for the synthesis of the proteins. So, it is actually and this is going to be a 5 prime UTR and the cap right. So, it is going to be a 5 prime cap and this remember that when we were talking about the post transcriptional modification we said that how the cap is been formed and how the poly A tail or the poly A dilation is happening.

So, messenger RNA has 3 reading frame out of which only one code for the desired protein. If it in the sequence of base there is no stop codon to interrupt the translation that met the entire peptide chain and that is called as the open reading frame. So, this is actually a coding sequence it is also going to be called as the open reading frame. Then let us talk about the this part right. So, it says that you are going to have the 3 reading frames.

So, when it says the 3 reading frame it actually is actually talking about the genetic codes right. And so, before we get into the discussion about the reading frames let us first discuss about the genetic code. So, genetic code so, messenger RNA is the random sequence of nucleotide differentiated by the base attached to them which are uracil, adenine, cytosine and guanine 3 nucleotide together code for a specific amino acid and these are called as the codons. So, you are actually going to have a codon where you are. So, it is actually going to be called as codon or I will say code right.

So, this code is made up of the 3 nucleotides right. So, for example, AUG ok and it always been 3. So, it has really going to be read as the 3 codon just like you know you are actually getting the OTP right. So, OTP is also 4 letter words right it is actually goes and give you the. So, when you do a you know banking transactions bank normally send you an OTP right.

That OTP is having 4 codes right 4 letter codes right. So, normally it is digits actually and when you enter that OTP the bank understand that it is actually the authentic person it is actually you who is doing the transactions. Similarly you are actually going to have the codons or the codes what is present on to the messenger RNA and all these codes are unique for their amino acids for example, the AUG is for methionine ok. So, similarly we have the codon for the other kinds of the amino acids. So, you see that so, codon has the 3 letters right first letter, second letter and third letter.

So, what you see here is the codon library. So, it is going to have the first letter, second letter and the third letter. So, if you go with the first, second and third what you see here is that you are going to have the multiple types of combinations and all these codons has to fulfill the criteria that it is actually going to be unique for that particular amino acid. It should be there should be no ambiguity right it should not be for other amino acids right. And the first question comes why there is a 3 nucleotide why not 4 just like in the OTP right you are getting the 4 amino acids.

So, you are getting the code which is of 4 digits right why not it should be 4 why not it should be 2 why not it should be 1 right. So, the people have done the experiments and people have mathematically also done the calculations and then they found that it is actually the 3 amino nucleotides which is actually been the optimal to provide the specificity and as well as the sufficient number. So, that you can be able to get the codon for all the 20 amino acids right. So, the genetic code is a triplet code right and it is called or scored it is known that we have only 4 type of nucleotide that make the whole genome. It is also known that the each codon consists of the 3 nucleotides which means that there are 4 to the power 3 that is the 64 possible amino acids.

However, since there are only 20 amino acids, it is obvious that more than one codon is used for a single amino acid. So, this is called the Bobel hypothesis. But there are 64 codons and only 20 amino acids. This is just the mathematical way of explaining that why there is a 3:1 ratio. Nucleotides are present in the codon, but people have done the experiments. So, what we have done is they are actually taken the codon for example, you can take like this you can make a poly A amino acid poly A messenger RNA for example, right.

So, if I use this right ok. So, this one is having the 3 3 3 3 right. So, this is going to be 3 prime this is going to be 5 prime like. Now, if you translate this under the in vitro translation system you are going to get a protein which is of the 4 amino acids right. So, it will be 4 amino acids. So, this is going to be the first amino acid this is going to be the second amino acid this is going to be third amino acid.

So, you are going to get a polypeptide which has the 4 amino acids. So, it is going to be a 4 amino acid polypeptide. Now, imagine that I am going to make a mutation ok. So, I am going to erase this right. So, for example, I have prepared another messenger RNA and I have removed one of the amino acids.

So, I have removed one of the nucleotides. Now, how many nucleotides do I have? I have 1 2 3 4 5 6 7 8 9 and 11 earlier I was having the 12. Now, how many amino acids am I going to get? I am going to get the 3 amino acid right. Now, again if I remove one more suppose I remove one more then I am going to have the 10 even then I am getting the 3 if I remove one more then also I am getting the 3, but if I remove one more right I am going to get the 2 amino acids because then this one also is not going to code. So, same similar these kind of you know where people have synthesized the small RNA species in the you know the in vitro translations and so on and by doing so, they could they could be able to identify the code for the different amino acids. So, for example, if you do a you are what you are going to code is actually the what you are going to get the code for the for example, you are going to get a code for lysine ok.

So, this is the genetic code for the lysine. So, this means you are going to get lysine lysine like that and that you can easily identify from the polypeptides right you can easily digest that polypeptide and it is actually going to tell you that this is actually the lysine. Same is true for other amino acids for example, I can make a messenger RNA with U U U right and I can say that oh it is actually coding for phenylalanine. Similarly, I can make the another amino acid like I can just make this kind of system and I can do a post you know in vitro transcriptions and then I will say ok U U A is coding for the lysine. But that is how they have actually come up with this kind of library which is called as the genetic code library ok or table ok genetic lab where this is the first letter this is second letter and this is third letter and you can easily

be able to know that ok if it is a U U U it is phenylalanine if it is C U U it is leucine and you see we have the 64 codons right. So, 64 codons for 20 amino acid that means, some of the amino acid will have more and more than one amino acids.

So, out of 64 the 3 amino acid 3 codon is for the stop codon and the 61 is for the amino acids right. So, what you see here is the 3 codons which are for the stop codon like the U A A U A G and U G A. So, these codons will not going to code for I mean any amino acids and that is why they are actually going to stop the synthesis of the translation this is very important to understand and it actually required when you are when we are going to discuss about the transformations. And so, this is this is very very you know detailed and remember that these codons are degenerate they are actually been present in all the organisms. So, same codon is being coded with in for a particular amino acid in all the organisms except there are few exceptions there are exceptions that some of the codons which are coding for methionine in the bacterial system, but they may they may be coding for something else in the eukaryotic system.

So, those kind of exceptions are there, but more or less more most majority of these codons are coding for the all coding for the same amino acid in all the organisms. So, these are the some of the properties of the genetic code. The genetic code is triplet and this means it is actually going to made up of the 3 nucleotides each codon coding sequence has a start and a stop codon to initiate and terminate the translation. Usually start codon is AUG right AUG which is for methionine right and which code for the methionine and stop codons are UA, UAG and UGA in some cases the starting codons are GUC or the UUC. So, these are the exceptions this are in the some of the bacterial system you are actually going to have the start codon as GUG or UUG.

We are going to discuss in detail why it is always methionine, but that you will understand once they are going to discuss about the initiation and other steps. The codon is the code is unambiguous which suggests that the code is for only one amino acid. That means, they are not going to get confused like. So, it is not like for methionine you have the same codon and the same codon is coding for the leucine also. Actually unambiguous it is only specific for that particular amino acids.

There is no gap and there is no comma in the codon. So, it is actually AUG there is no difference between these two rights there is no gap between or there is no comma actually. The code is degenerate this means that one amino acid has more than one codon for example, the phenylalanine is specific for two codons that is a UUU and UUC that this is the phenylalanine UUU and UUC. So, this is actually degenerate. So, you can actually have the flexibility of using a tRNA which contains the UUU or you can use the tRNA which is for a UUC. And it actually provides the flexibility into the system some

organism may use the UUU very extensively and some organism may use the UUC and that is why it may actually help the organism to adopt according to the environment or according to the situation.

Only tryptophan and the methionine are coded by the single codon. Then the codon is non-overlapping for example, a code has AUG, CUG, GGU, GAU, UUU, GUA then codes will be AUG right. So, you go by 3 3 like. So, GU CUG, GGU and so on, but not like AUG, UGC, GCU and so on. So, this is what is actually going to say that there is only one reading frame which is possible right.

So, there are suppose I will explain you this in that. So, for example, you have AUG and CUG. So, I will just explain with two codon actually. So, what it says that you actually going to read this as first codon and you are going to read this as a second codon, but you are not going to read this as a overlap. For example, you cannot read like this you cannot read like you can read like AUG and CUG, but you cannot read like this you cannot read like this because then it is actually going to be a problem right. So, if you are actually going to have the codons it can actually be able to and if you go by this you can be able to have the 6 different types of reading frames ok, but ideally we only have the one correct reading frame which actually provides the correct amino acids.

Genetic code is universal which suggests that the genetic code and its meaning is common for all life forms. However, there are some exceptions in this rule. For example, UGA is a stop codon, but it codes for tryptophan in the mycoplasma, spiroplasma and the mitochondria of several species. Similarly, CUG codes for the leucine in general, but in yeast mitochondria it codes for the threonine.

So, these are the some of the exceptions. Now, let us move on to the next structure and the next structure is the transfer RNA ok. So, the transfer RNA is a clover leaf structure in the 2 dimensional and L shaped structure in the 3 dimensional. Transfer RNA is a 73 to 94 ribose nucleotides in length. A transfer RNA molecule consists of 5 prime phosphate terminal and acceptor arm and ends in the CCA terminal at 3 prime end. D loops which often contains the dihydro uridine, anticodon loops and the t arm which has the t psi c where psi is a pseudo uridine.

CCA sequence is important as it is important for the recognition of tRNA and also site for attachment of the amino acid. So, this is the clover leaf model of the tRNA right. So, where you are going to have the 5 prime end and from the 5 prime end you are actually going to have the 5 different arms right. You are going to have the d arm, you are going to have the anticodon arms, you are going to have the di-cis group and you are going to have the CCA end.

And then you are also going to have the extra arm. This extra arm is being used for adjusting the extra nucleotides. So, it is actually used for adjusting the extra nucleotides. So, that you can always have the anticodon in a particular orientation, you also should have the CCA end in a particular orientation. These are the these 2 arms are actually containing the specific nucleotides for example, the d loop is having the dihydro uridine as a nucleotides. Similarly, t psi c has the pseudo uridine as one of the modified nucleotides what is present.

And then this is a very important codon because this is a anticodon loop which is actually going to recognize the codon onto the messenger RNA. So, it is actually going to recognize the codon what is present onto the messenger RNA. For example, if you have AUG, this AUG what is present onto the messenger RNA is going to be recognized by the anticodon what is present onto the tRNA. Similarly, this is the amino acid arm or CCA arm. So, this is the CCA arm where your amino acid is actually going to be attached and that is how this is actually going to participate into the condensation reaction.

So, this part is going to you know participate into the condensation reaction and this part is actually going to read the message onto the messenger RNA and that is how they are actually going to bring the correct tRNA into the actions. So, each tRNA is specific to the amino acid that it carries in the CCA arm. So, there are 30 to 45 different tRNAs in prokaryotes and the 50 different types in eukaryotes which suggest there is a more than one tRNA for a single amino. So, tRNA would be as per the genetic code. So, you are going to have the different types of anticodons for the different amino acids.

So, remember that we have the different types of genetic code for a specific amino acid or for a particular amino acid. Similarly, you are going to have the different types of tRNAs. So, you are going to have the 50 different types of tRNAs in the pro eukaryotic system and 30 to 45 different types of tRNAs in the prokaryotic system. For example, for glycine there are two tRNAs which are represented as tRNA gly 1 and is also called as tRNA gly 2 right.

So, these both are actually going to carry the glycine here. So, it is going to carry the glycine here, but the anticodon are probably different. Now, let us move on to the third machinery and the third machinery is called as the ribosome. So, structure of the translational machinery. So, ribosomes are the ribonucleotide proteins. So, these are the multimeric protein complexes where you are actually going to have the ribosomal RNA and the proteins.

Each ribosome is made up of the two subunits large subunit and the small subunit. In prokaryotes mitochondria and chloroplast of the prokaryotes chloroplast there is a 70S ribosome which is composed of 50S and 30S subunit. In equalize the 30S subunit consists of 16S ribosomal RNA, 21 ribosomal RNA and 50S ribosomal subunit and 50S subunit contains 23S ribosomal RNA, 5S ribosomal RNA and the other 31 different types of proteins. In eukaryote there is a 80S ribosome which is consists of 60S and the 40S ribosomal subunit. So, in eukaryote there is a 80S ribosome which is consists of 60S and 40S ribosomal subunit, 60S subunit consists of 20S RNA, the small ribosomal RNA and

the

5.

8 ribosomal RNA and approximately 50 different types of proteins. The 40S ribosomal subunit consists of 18S ribosomal RNA and the 33 different types of proteins. So, this is what and it is actually going to have the 3 different types of activities is going to have the aminoacyl site, it is going to have the peptidyl site and it is going to have the exit sites. And one side it is going to have the messenger RNA binding site and on the other side it is actually going to have the tRNA docking site. So, the ribosomes are ribonucleoprotein that contain the RNA and protein and each ribosome is made up of the 2 subunits. So, it has actually going to so, 70S ribosome has 3 RNA tRNA binding sites has the P site, it has A site and has E site.

The P site is the peptidyl tRNA binding site, A site is the aminoacyl tRNA binding site and E site is the deacetylated tRNA site all the exit sites. So, it is going to have the 3 sites going to have the P site going to have the A site, it is going to have the E site. So, A site from where the RNA the tRNA is going to enter into the protein synthesis then it will enter it then it will move to the P site and it is actually going to catalyze the reaction of condensation and then from here it will either move to this site or it will actually move to this site depending upon what kind of codons are present at the bottom. If it is codon is next codon then it will enter here and then it will again take up the another amino acid, but if it is a stop codon then it will enter here and it will be get exited from this site ok. That anyway we are going to discuss in detail when we are going to discuss about the mechanism of the translation.

So, mechanism of translation so, it is going to have the 3 important events like the activation, initiation, elongation, terminations and we are also going to discuss about the post translational modifications. So, mechanism of translation so, first event is the activation of amino acids. During this process the amino acids are attached to the tRNA in the presence of the enzyme aminoacyl-tranvasanthase this enzyme activates amino acids by attaching covalently to the tRNA when the tRNA get charged it named it named as the amino T aminoacyl tRNA aminoacyl tRNA. During this process the amino acids are attached to the tRNA with a high energy bond so, that the it activates the amino

acids. So, you are going to have the amino acid you are going to have tRNA you are going to have ATP then you are going to have an enzyme which is called as aminoacyl-tRNA synthetase and it is actually going to produce the aminoacyl tRNA and it is going to produce the AMP and Pi which is actually going to be produced from the ATP.

This is actually very high energy system where your amino acids are going to be attached on to this right and it is actually going to participate into the condensation reactions. In eukaryotes in eubacteria first amino acid in the polypeptide chain is N-formylmethionine which is specific to 3 codons as AUG, GUG and UUG. Whenever these codes are present in the primary point they code for the N-formylmethionine but they are present in between the codon sequence they code for the methionine and valine respectively. How does this happen? This is because of the difference in the initiation tRNA and the one used in between the process of translation. Initiation tRNA has a unique feature that distinguishes from the elongating tRNA in the bacteria.

So, you are going to have the amino acids tRNA and ATP and it is going to have synthesis of the amino acid tRNA. The only difference is that you are going to have the different initiation tRNA compared to the tRNA what is being used for that during the elongation steps. Now let us talk about the mechanism. So, mechanism has the 3 steps going to have the initiation, elongation and terminations. So, in the first step the small subunit of ribosome binds to the messenger RNA such that the initiation codon lies in the P-side.

This gets possibly due to the activity of IF3 it basically prevents the untimely reassociation of the large and small subunit of the ribosomes. Moreover, it promises accuracy of initiation site selection. There is a ribosomal binding site which consists of the Shine-Dalgarno sequences and the initiation codon. This Shine-Dalgarno sequence which is 5' prime A and this is located 10 base pair upstream of the initiation codon is complementary to the region near 3' prime end of the 16S rRNA a component of the small subunit of the ribosomes. So, this is the Shine-Dalgarno sequence what is going to be present on to the initial on to the messenger RNA which is upstream to the start codon.

And then in the next step the initiation tRNA carrying the N-formylmethionine enters the P-side and binds to the messenger RNA via its anticodon loop. So, in this is the initiation codon. So, it is actually going to have the binding and the assembly of the ribosomal machinery in such a way that this actually is going to bind to the P-side. And then once it binds to the P-side and then the IF2 is responsible for this activity. So, with the help of the initiation factor 2 it is actually going to allow the binding of the tRNA on to the P-side.

So, on the P-side the tRNA is actually going to recognize this initiation RNA and initiation t-methan-codon and that is how it is actually going to bind with the initiation amino acid. So, that is the 5-formylmethanine. It directs the initiation tRNA to its correct position in the initiation complex. It is also exhibit ribosome dependent GTPase activity.

Once the GTPase hydrolyze then the 50S ribosome and join to form the complete ribosomes. Finally, when the large subunit also join the complex it forms the complete P-side and the A-side. Second charge tRNA enters into the A-side this tRNA as per the rule has the anticodon corresponding to the codon into the messenger RNA. So, once the tRNA the initiation tRNA is going to be placed very nicely then the 50S ribosome subunit is going to enter. It is going to replace the initiation factor 1, 3 and 2 and it is going to hydrolyze the GTP as well. And then the large subunit will also going to bind in such a way that it is actually going to make the complete P-side, complete A-side and so on.

And then the P-side is actually going to take up the next tRNA that next tRNA would be the specific for the next codon. Now, let us move on to the elongation steps. So, it is a cyclic process elongation process starts from the formation of first peptide bond to addition of last amino acids. This amino acid added to the chain 1 at the time to nascent polypeptide chain. Addition of amino acid is very rapid process the peptide sequence is in order of codon and anticodon in the messenger RNA rate of elongation is nearly 15 amino acids per second.

So, there are some requirement regarding to the elongations that the messenger RNA and 70S ribosomes, aminoacyl tRNA and the elongation factors. So, you have the 3 different types of elongation factor elongation factor TU, elongation factor TS and elongation factor G. So, elongation factor TU it is a G protein which binds to the aminoacyl tRNA and direct it to the correct position and the ribosome A-side. Its main function is to regenerate the EF-TU and the hydrolysis of the GTP and then EFG it also has a G protein which mediate the translocations.

So, elongation is carried out by the ribosome in 3 steps. Remember that we have the 3 activities we have the P-side, we have A-side and we have E-side. And we already have the amino acid. So, in the initiation side we already have amino acid in the A-side in the P-side and it only moved. So, we actually have a tRNA which contains the another amino acid in the the A-side.

So, first it is actually going to have the decoding. So, it is a codon directed binding during the process of ribosome select and bind to the incoming aminoacyl tRNA at a site

whose anticodon is complementary to the codon of the messenger RNA. Decoding region of 16 as ribosome confirm the proper base pairing between the codon and the anticodon. So, on the A-side there will be a codon and anticodon interactions. So, this is going to be messenger RNA what is running right and this is going to be anticodon what is present onto the tRNA and that is how this interaction will ensure that the correct amino acid and correct tRNA is actually going to bind into the A-side.

Then we are going to have the peptide bond formation. So, in this process the peptidyl group of P-side of tRNA is transferred onto the amino acid group in the peptide bond. So, this amino acid is actually going to be shifted onto this and that is how you are going to have a peptide bond which is going to be formed and this I mean this is going to be a free right. And then there will be a translocation. So, this there will be a translocation this is this whole complex is going to be translocated onto this side and that is how whatever is this dipeptide is going to be present onto the P-side.

And then again the third second nucleotide is going to be present here. So, in this case the tRNA of A-side is transferred onto a P-side to make a space for the next amino acid tRNA at A-side and the A-side of tRNA is shifted at the E-side and this shift is also coupled with the ribosome movement along with the messenger RNA. So, then so process of chain elongation in ribosomes. So, EF-Tu promote the entry of amino acid transfer A into the A-side of the ribosomes. First EF-Tu binds to the GTP and it activated the EF-Tu GTP complex which binds the tRNA when codon and anticodon base pairing stabilizes then hydrolysis of GTP occurs which converts into GDP and PI which helps in the binding of amino acid tRNA to A-side and after this EF-Tu is released. EF-Tu is catalyzed the release of the EF-Tu GTP from the ribosome and regenerate the EF-Tu GTP its main work is to recycle the EF-Tu.

So, this is what is actually going to happen in the elongation steps. And then you are going to have the peptidyl transferase. So, it is a peptide bond formation step which with which the amino acid of the peptide bonds are linked to the tRNA in A-side and a carboxyl end of the peptide chain uncoupled from the tRNA by in the P-side. This reaction is carried out by an enzyme which is called as peptidyl transferase. Peptidyl transferase is an enzyme which is associated with the 23S ribosomal RNA of 50S ribosomal

Peptide bond formation involves the O₂N migration and the conversion of ester into the amide bond. So, this is what exactly going to happen in the P-side you are actually going to have the amino acid 1 and from here it is actually going to be shifted onto the amino group of the second amino acids and that is how you are going to have this dimer which actually going to have the peptide bond. So, this is actually going to be a peptide bond

what is going to be formed. So, three things are necessary for the translocations right deacetylated RNA moves from the P-side peptidyl transferase move from A to P-side and the ribosome should move on to the messenger RNA 1 codon so that the next codon can come at the A-side. Translocation steps carried out by the EFG factors during translation acceptor end of the both tRNA of A and P-side are interacting with the peptidyl transferase center of 23S ribosomal RNA of 50 subunit. In trans location A and P tRNA transfer to the P and E-side respectively as ribosome move 3 nucleotide along messenger RNA chain in 5 prime to 3 prime direction.

During trans location step the GTP is converted into GDP and the uncharged cRNA released from the P-side to E-side and newly formed a beta transfer from A to A-side. In elongation process is nearly same in both the prokaryotes as well as the eukaryotes. This is what exactly going to happen this is the assembly of the system and this is going to be A-side, P-side and E-side. So, A-side is always going to welcome the new amino acid or new amino acid transferase. Whereas in this side you are going to have the peptide bond formations and once the peptide bond is been formed then it is actually going to be shifted on this side.

Whereas the old tRNA which does not contain amino acid is actually going to be shifted onto the E-side and from here it is actually going to be go into the cytosol. Because so this tRNA is going to be rejected into the cytosol and then again it will participate into the another round of the amino acetylations with the help of the amino acids and again it will be ready for supplying them in the amino acids into the A-side. And then if we have the terminations so termination of translation occur due to the stop codon there are three stop codons UA, UAG and UGA. Out of these three when one of the stop codon appears in the A-side of the ribosome it causes the termination because there is no tRNA present corresponding to these codons.

So, tRNA is not binding codon and cause the terminations. During the termination process release factors are involved when the UA or UG is in A-side RF1 binds to the ribosome when UA or UGA is in the A-side RF2 binds to the ribosome. RF3 is a type of GTPase which maintains the function to catalyze the release process to GTP binding and hydrolysis. So, in the terminations release factor RF1 or release factor RF2 binds to the ribosome nearly at the A-side then the polypeptide chains are released from the ribosome by the peptidyl transferase complex. So, peptidyl transferase complex transfers the carboxyl terminal residue of the peptide chain from the tRNA of P-side to the water molecule.

Now the release factor RF and GTP released and its tRNA is also freed. Now the

ribosome assembly is unstable due to the presence of the initiation factor IF3 and IF1 and ribosome recycling factors as a result the ribosome 70S ribosome disrupt into the 30S and 50S ribosome and prepared for the initiation. So, this is exactly happened once you have the complete synthesis and you are actually going to reach to a stop codon then on the stop codon the release factors are actually going to participate and they are actually going to release the newly synthesized proteins from the newly synthesized protein from the ribosome. And once the protein is been released then the ribosomal large subunit and the small subunit is also going to assemble going to be disassembled and they are actually going to participate into the initiation with the another codons and another set of the translational machinery. So, this is what we have discussed so far we have discussed about the translational machinery we discuss about the structure of the tRNA we have discussed about the structures of the messenger RNA we discuss about the structure of the ribosomal RNA or ribosomes and then we also discuss about the different events. So, we discuss about the genetic codes we have discussed about the initiation, the elongation and terminations and we have also discuss about the different types of processes what is happening into the prokaryotic system.

In our subsequent lecture we are going to discuss about the eukaryotic translations and how it is different from the prokaryotic translations. So, with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss about the eukaryotic translation and how it is different from the prokaryotic system. Thank you. .

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 07
Translation
Lecture-31 Translation in Eukaryotic System

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering at IIT Guwahati and what we were discussing in the course molecular biology. We are discussing about the cell biology, we discuss about the cellular processes, we discuss about the different types of organelles and then we have also discuss about the biomolecules their function and then followed by we have also discuss about the genetic material. So, we discuss about genetic material in the prokaryotes, eukaryotes and how the genetic material is been packed into these two different types of organisms and subsequent to that we were also discussing about the central dogma of molecular biology. So, within the central dogma of molecular biology so far what we have discuss we have discuss about the replication, transcription and in this current module we are discussing about the translations. So, as we said that the central dogma of molecular biology is a very important phenomena in which the three different types of activities such as the production of origin or the synthesis of the new DNA from the pre-existing DNA is been done by a process known as replications whereas, the RNA is been produced from the DNA with a process known as transcription and in today's and in this current module we are discussing about the translation where the information what is given on the RNA is been used to synthesize the protein molecules.

If you recall in our previous lecture we have discuss about the prokaryotic replication. So, we have discuss about the different types of machinery what is required for prokaryotic applications so we discuss about the ribosomes tRNA messenger RNA and so on. And then we have also discuss about how the different types of events are happening such as initiation, elongation and terminations in the case of the prokaryotic system. In today's lecture we are going to discuss about the eukaryotic system.

So, before we get into the detail of what is the major difference of the translation in the prokaryotic versus eukaryotic we would like to you know give you a very small description about the machinery what is required for the translations and then we are actually going to tell you in detail about the different processes what is happening in the case of the eukaryotic translations. So, eukaryotic translation is a process or the translation is the process of messenger RNA coded protein synthesis this is a universal process that occurs both in the prokaryotes and the eukaryotes. Members of the prokaryotes or the eukaryotes use the information what is given in the messenger RNA

which comes from the DNA by transcription to synthesize the protein with the ribosome as a machinery. So, protein form a variety of critical functions such as enzymes, structural proteins or hormones and therefore, they are crucial for the biological component and that is why we said that the central dogma of molecular biology is very crucial because it explains how the different events are connected to each other and why they are so much crucial because once you have a requirement of a particular hormone you are actually going to activate the transcription of that particular gene and then you are actually going to activate the translational machinery. So, that the ribosomes will go and sit on that particular RNA and it will going to give you the proteins.

Protein biosynthesis has a key role in the disease as changes and occurs errors in this process through underlining DNA mutations or protein misfolding are often the underlining cause of the disease. And protein machinery is very crucial because it can give you a misfolded protein, it can give you a protein which may have mutations, it may give you the protein which may not be useful for or it may not be optimal for you know doing the its natural function and because of these things it may actually lead to the development of the disease. So, the process by which the sequence of nucleotide in a messenger RNA molecule direct the membrane incorporation of the amino acid into the protein is called as the translation. So, this is all we have already discussed but I thought when we were discussing about the eukaryotic translation we should also briefly discuss about these aspects as well so that it will refresh your memories so that it will be easy for you to follow the follow the content actually. Now, as far as the machinery is concerned the machinery required for translating the language of the messenger RNA into language of protein is composed of the four primary components so it requires the messenger RNA right so that is the component number one and that is the most crucial component.

Then you requires the tRNAs so you require the tRNA so that it can actually be able to read the anticodons with the help of the codon versus anticodon interrecon and the other side it also can supply the specific amino acids then it also requires the ribosomes and it also requires the different types of proteins and as well as the enzymes so that it can actually be able to perform the different types of activities such as aminoacyl tRNA oscillations and all that and peptidyltransferase. So enzymes are two there are two crucial enzymes you have you require the aminoacyl tRNA synthetase and you also require the peptidyltransferase as well as the protein is concerned you will also require different types of translational factors we have discussed many of these translational factors when we were talking about the prokaryotic translations. So let us start with the first component and the first component is the messenger RNA so messenger RNA is a single standard RNA molecule that is complementary to the one of the strand of a gene during the protein synthesis ribosomes moves along the messenger RNA read its waste sequences and uses the genetic code to translate the each codon into a corresponding

amino acid so this is the eukaryotic messenger RNA where you are going to have the 5 prime UTR regions you are going to have the coding sequence and then you are also going to have the 3 prime UTR region and then you are going to have the poly A tail and that majority of the 5 prime UTR region or the 3 prime UTR region is actually the regulatory regions where the many of these regulatory proteins are going to bind and that is how they are actually going to regulate the translation within the coding sequence you are going to have the starting codon and you are also going to have the stop codon. So in the case of eukaryotic system you are going to have the start codon as AUG and which is actually going to code for methionine whereas there are 3 stop codons UA, UAG and UGA and these are the 3 stop codon which are also been the stop codon into the prokaryotic system. Then we have the tRNAs I am going to I am going through very fast with these content because the already we have discussed this is the you know the nucleotide sequence of the tRNA where you have this d arm right or then you are going to have the anticodon arm then you are going to have t psi c arm and this is actually going to be start from 5 prime end you are going to have d arm anticodon arm loop t psi c loops and the cca end and the cca end you know that it is actually going to bind the amino acid on one side whereas the anticodon loop is going to have the anticodon which is actually going to recognize the codon on to the on the messenger RNA and that is how it is actually going to serve the dual purposes. It is going to serve it is going to first identify the messenger RNA on one side and it is going to bring the corresponding amino acid from the 3 prime ends. So the primary sequence primary structure of tRNA is a linear sequence of the nucleotides secondary structure is called as the clover leaf models and the tertiary structure is called as the 3D structure of tRNA or the L shape or the helix packing. So these are the some of the different names what we are using the tRNA is also called as transfer tRNA transfer RNA is a type of RNA molecule that help to decode a messenger RNA sequence into a protein and it is made up of a single standard poly nucleotide chain. It function at a specific site in the ribosome during translation which is a process that is inside the protein from a messenger RNA molecules.

Proteins are built up from the smaller units called as the amino acids which are specified by the 3 nucleotide messenger RNA sequence called as codon. All these we have already discussed when we were discussing in detail when we were discussing about the prokaryotic system. So each codon represent a particular amino acid and each codon is recognized by a specific tRNA. They are adapted between the codon and the amino acids each tRNA has its corresponding amino acid attached to its 3 prime end and the tRNA is named as sRNA or the soluble or the pneumatic RNA and the adapter RNA. tRNA the 10 to 15% of the total cellular RNA which actually includes the messenger

RNA or ribosomal RNA and the tRNA.

So out of these 3 RNA species total of the 10 to 15% is the tRNA species 74 to 95 nucleotides are present in each tRNA molecules then it has the 3.8 sedimentation coefficient and the molecular weight of the tRNA is between the 25,000 to 30,000 Dalton. The structure of tRNA can be decomposed into the primary structures secondary structures and the tertiary structures. Secondary structure is also called as the cloverleaf model whereas the tertiary structure is called as L-shaped structure. In addition to the usual nucleotide bases it also tRNA contains a number of unusual bases such as ionosine, pseudo uracil and dihydro uridine and these are the amino acids modified nucleotides which are been by methylation.

So for example the ionosine is going to be produced from the adenine, pseudo uridine uracil is from produced from the uracil and pseudo uridine is produced from the uridine. The other unusual amino acid found is hypoxanthine, thymine and methyl guanine. So this is the structure of the tRNA this is the cloverleaf model of the tRNA where you have the 3 prime end which is also called as the CCA end then you all going to have T-SciC loop, D-loop, and decodone loop and all that. So 5 prime terminal is a phosphate group and then you are going to have the acceptor stem it is the 7 to 9 base pair stem by the base pairing of 5 prime nucleotide with the 3 prime nucleotide then you have a CCA tail it is a cytosol, cytosine amino acid sequence at the 3 prime end of the tRNA molecule the amino acid loaded onto tRNA by the amino acid tRNA synthesis to form the amino acid tRNA is covalently attached to the 3 prime hydroxyl group onto the T-R CCA tail. D-loop it is a 4 to 6 base pair stem ending the loop that often contains the dihydrouridine then you have anticodon loop it is a 5 base pair stem which loop contains the anticodon and anticodon is going to recognize the genetic code or the codon what is present onto the messenger RNA and then you have a tR it is a 4 to 5 base pair stem containing sequence of T-SciC where psi is the pseudouridine a modified uridine actually.

Let us talk about then the ribosomes so ribosome is actually the real machinery of the protein synthesis which is different from the prokaryotic system versus eukaryotic system in the prokaryotic system you have the 70S ribosome whereas in the eukaryotic system you have a 80S ribosomes. So eukaryotic ribosomes are larger they are 80S ribosomes and a more complex than the prokaryotic ribosome which are 70S. Ribosome exists normally as a separate subunit that are composed of the proteins and the ribosomal RNA the subunits come together to form a ribosome when they bind to a messenger RNA near its 5 prime end on binding to the messenger RNA the ribosome read the nucleotide sequence from the 5 prime to 3 prime directions synthesizing the corresponding protein from the amino acid in a N terminal to C terminal directions. Ribosomes are located in the cytoplasm either freely floated or the associated with the

endoplasmic reticulum. They serve to synthesize the proteins the ribosomes are ribonucleoprotein particles to which the multiple ribosome proteins are bound the sequence and the structure or ribosomal components are conserved in all kingdom under underlining the common origin of the translational operators.

The ribosome provide the platform for proper assembly of the messenger RNA tRNA and the protein factors it consists of a small and the large subunits. So this is the structure of these large subunit and this is a structure of the small subunit and it has a three different types of binding site that has E site it has a P side and has a A side and the mammalian ribosome which is a 80S ribosome is consist of the two subunit this is a large subunit and this is a small subunit and a large subunit is composed of the different types of RNA species such as 28 ribosomal RNA 5.8 S ribosomal RNA and 5S ribosomal RNA and then it also contains the 49 different types of proteins. So all these 49 different types of proteins when they come together along with the 28 S ribosomal RNA 5.8 S ribosomal RNA and 5S ribosomal RNA that is actually going to give you large subunit which is going to have segmentation coefficient as 60S then you also have the small subunit which is a 40S subunits and that contains the 18 S ribosomal RNA and the 33 different types of proteins and when they come together they are actually going to make the ribonucleoproteins and they are actually going to make the small subunit which is going to be 40S and together they were actually going to when they will combine together they are actually going to give you the 80S complete ribosomal particle which is going to participate into the protein synthesis.

Now as far as the binding site is concerned the t are actually going to have the three binding sites for tRNAs they are going to have A site, P site and E site. A site is the site where the new amino acid or the new tRNA is actually going to enter the P site is the site where the peptide bond is going to be formed and the E site is the exit site from where the uncharged tRNA is actually going to get removed. So all three sites are formed by the ribosomal molecules into the ribosomes during the elongation the incoming RNA molecule binds to the A site the P site is where the tRNA linked to the growing polypeptide chain is bound and the E site is the bonding site for the undoded tRNA prior to its release from the ribosomes. Now as far as the translation in the eukaryotic is concerned they are actually going to have the three important events initiation, elongation and terminations. So initiation sets the stage for the polypeptide synthesis so it actually going to assemble all the protein components it is going to assemble the ribosomes it is going to have all those events so that the it is going to you know bring the raw material actually and then it is going to enter into the second phase which is the elongation.

So that causes the sequential addition of the amino acids to the polypeptide chain as

determined by the codons what are present on to the messenger RNA and then you are going to have the termination so this bring the polypeptide synthesis to a halt because once it reaches to a stop codon then it is actually going to terminate. So let us first start with the initiation so initiation if you recall we have discussed about the initiation in the prokaryotic system. Now when we are going to discuss about the initiation into the eukaryotic system we just want to first understand what is the difference between the two different types of events before we get into the detail of the initiation into the eukaryotic system. There are significant differences between the initiation of the initiation stage of the prokaryotes and the eukaryotes. In eukaryotes there are only one start codon for the eukaryotes such as AUG and it is codes for the methionine it does not code for the N-formyl methionine.

So eukaryotic cell need more initiation factor than the prokaryotic system. For example the eukaryotic cell requires the 12 different types of initiation factor whereas the prokaryotic system requires the lesser number of initiation factors. In prokaryotes the presence of association of messenger RNA with the small subunit is more complex than the eukaryotes. 40S subunit identify the 5 prime methylated cap of the messenger RNA and there is a scanning process involved whereas the initiation codon is recognized. This recognition is added by the ATP dependent helicases that hydrolyze the ATP.

This recognition of initiation codon is also been aided by the COSAC sequences and COSAC sequence are very much same as what we have seen in the role of the Scheinder-Ganno sequences into the prokaryotes. So let us discuss about the initiation. So the initiation of the translation in the eukaryotes involve the many initiation factors or the EFIs and it is divided into the four stage. In the stage 1 the ribosome is going to be dissociate. In the stage 2 the complex of 43 pre initiation complex is going to be formed and then this 43S initiation complex is going to be converted into the 40S initiation complex and then ultimately it is actually going to have the formation of 80S initiation complex.

This means the ribosome is going to be fully assembled onto the messenger RNA and then it will enter into the elongation phase and that is how it is actually going to start the initial elongations. So during this particular phase it requires the many type of accessory proteins and initiation factors for performing the different types of functions and as I said in the formation of these kind of pre initiation complexes. So it will require the different types of initiation factors. So some of the initiation factors are called as core initiation factors whereas the other factors are called as the accessory initiation factors. So these are the core initiation factors.

So you have the EF1 and EF2 and EF1A and that enhance the pre initiation complex

formation helps in ribosome scanning, assure the steadiness of the AUG selection, prevents the premature hydrolysis of the EF2. Then we have the elongation initiation factor 2 and that assists the binding of the methionine tRNA met to the 40S ribosome by forming a ternary complex of the initiation eukaryotic initiation factor 2 GTP and met tRNA. Then you have the eukaryotic initiation factor 2 and eukaryotic initiation factor 3. So the first initiation factor binding to the 40S subunit and promote the further steps having the you know the gonadine nucleotide exchange factor activity. So and then we have the eukaryotic initiation factor 4A that contains the 2 domains.

So it has the dead box ATPS domain and the ATP dependent RNA helicase activity. Then we have the initiation factor 4B that is the co-factor of the initiation factor 4A and it enhances the helicase activity of the initiation factor 4A. Then we have the initiation factor 4E and that binds to the 5 prime cap of the messenger RNA that is the M7 GPPPG cap right. Then we have the initiation factor 4F that recruits the 40S to the 5 prime end of the messenger RNA. Then we have the initiation factor 4G and that scaffolds for the initiation factor 4E, initiation factor 4A, initiation factor 3, initiation PAB, slip 1 and messenger RNA and also participate in enhancing the helicase activity of the initiation factor

4A.

Then we have the initiation factor 4H that enhances the helicase activity of the initiation factor 4A and then we also have the initiation factor 5A and 5B that having the JTPS activity that hydrolyzes and promote the dissociation of the various initiation factors from the 40S and also leads to the association of the 60S subunit to form the ATS ribosomes. And then we have the auxiliary initiation factors. So these are called as DHX29, DED1 initiation factor 6, P97 and PAB. So DHX29 is having the helicase activity which contains the DED box functioning in the initiation step and also requires for the ribosomal scanning onto the messenger RNA. Then we have the DED1 and it is a homologous of DHX29 found in the *saccharomyces cerevisiae*.

Then we have the initiation factor 6 that binds to the 60S subunit and prevents it binding with the 40S subunit. Then we have P97 that is homologous to the C terminals of the initiation factor 4A and considered as the translational repressor under the normal cellular conditions. Then we have PAB it binds to the poly A tail of the messenger RNA also with the initiation factor 4A and initiation factor of RF3A. It promotes the circularization of the messenger RNA and it stimulates to the 40S subunit recruitments. So the first step is the RB ribosomal dissociation.

So ATS ribosome dissociate to form the 40S and the 60S subunits. The two initiation factor namely the initiation factor 3 and initiation factor 1A binds to the newly formed 40S subunit thereby block its re-association with the 60S subunits. This means this

ribosome is going to be first the ribosome where you know initially the ribosome would be involved into a protein synthesis and as soon as it will reach to the termination site then it will get dissociate you know dissociate and it is going to dissociate into the 40S and the 60S ribosomes. Once the 40S is been produced then it is actually going to be bind by the initiation factor 3 and initiation factor 1A and it is going to bind in such a way that it is going to block the association of these two because they will get recruited and they will be get associated but with the new messenger RNA and then only it is actually going to start the initiation of the new messenger RNA. So for initiating the new cycle it has to be dissociated from the older cycle.

So it is actually present on the some other messenger RNA and it reaches to the termination site. So at when it reaches to a termination site that the protein synthesis is going to stop but the ribosome has to be dissociated ribosome has to be broken apart so that you are going to have the 40S and the 60S subunits available and these 40S 60S subunit will assemble on to the new messenger RNA on which it is actually going to start the initiation and that is why this is actually the first step right and the second step is that the formation of 43 pre initiation complex. So a ternary complex containing the initiation tRNA that is the met tRNA and initiation factor 2 bounce to the GTP attached to the 40S subunit to form the 43 pre initiation complex. The presence of the initiation factor 3 and initiation factor 1A stabilizes this complex. So remember that these two factors are also important for you know blocking the attachment site so that it should not form the 80S ribosomes.

Then we have the formation of the 48 initiation complex. So the binding of messenger RNA to the 48 3 S initiation complex results in the formation of 48 initiation complex to the intermediate 43 S initiation complex. So initiation factor 4F complex is formed by the association of the initiation factor 4G initiation factor 4A and initiation factor 4E. The initiation factor 4F referred to as a cap binding protein binds to the cap of the messenger RNA and then the initiation factor 4A and the 4B binds to the messenger RNA and reduces its complex structure. This messenger RNA is then transferred to the 43 S complex for the appropriate association of the 43 pre initiation complex with the messenger RNA energy has to be supplied by the ATP.

The ribosomal initiation complex is scanned messenger RNA for the identification of the appropriate initiation complex that is the 5 prime AUG and then it is actually going to initiate the formation of the 80S ribosomes. So the next step is the formation of the 80S initiation complex. So 48 initiation complex binds to the 46 60 S ribosomal step unit to form the 80 S initiation complex. The binding involves the hydrolysis of GDP this is bound to the initiation factor 2 and this step is facilitated by the involvement of the initiation factor 5. As the 80 S complex is formed the initiation factor bound to the 40 S

initiation complex are released and the recycle.

So these are all the events what is being shown here that you are going to have the dissociation of the ribosomes. So in the first step you are going to have the dissociation of the ribosomes. So 40S and 60S which are present on to the previous some other messenger RNA where they have actually completed protein synthesis this is actually going to assemble going to dissociated and the 60 S is going to be get separated and the 40 S is going to get separated. So as soon as the 40 S is going to form it is actually going to bind the initiation factor 1 and initiation by 1A and the initiation factor 3 and then it is actually going to bind all these you know metionine metatRNA and mettRNA initiation factor 2 GDP and all that and that's how it is actually going to form the 43 initiation complex.

So that is going to be the second step. And once the 43 initiation complex is formed then it is actually going to bind the initiation factor 4E, 4G, 4A and 4B and after that it is actually going to be good enough to recruit the messenger RNA and once the messenger RNA is been recruited then the 43 messenger RNA complex is going to form on to the messenger RNA and then it is going to start scanning with the help of the ADP hydrolysis to know where the starting codon is right and then it is actually going to form the 49 S pre initiation complex and once it find the pre initiation complex 49 initiation complex and it found that there is a AUG right then it actually going to recruit the 60S ribosomes and 60S also going to assemble and that's how it is actually going to form the 80S initiation complex on to the initiation codons. Now it enters into the elongation phase. So ribosomes elongates the polypeptide chain by sequential addition of the amino acids. The amino acid sequence is determined by the order of codon into the specific messenger RNA elongation which is a cyclic process involve the certain elongation factor or EFs. Elongation can be divided into 3 steps binding of the aminoacyl tRNA into the A site peptide bond formation and the translocations.

This we have very detail in this we have discussed when we were talking about the prokaryotic system how the incoming aminoacyl tRNA is going to bind into the A site and then the existing peptide bond is actually going to form existing peptide chain is actually going to form the peptide bond and then it is actually going to be translocated on to the A site and then there will be a translocation so that newly formed peptide bond containing the peptide chain will get translocated and will bind the P site and then the tRNA what is present on to the P site will not have the any amino acid so it is going to be present on to the E site and from there it is actually going to get removed from the ribosomes. So this we are not going to discuss in detail. So the ATS initiation complex contain the met tRNA in the P site and the A site is free. Another aminoacyl tRNA is located into the A site this require the proper codon recognition on the messenger RNA

and involvement of the elongation factor 1A and the supply of energy by the GTP. The aminoacyl tRNA is placed in the A site elongation factor 1A and the GDP are recycled to bring another aminoacyl tRNA.

Then there will be a peptide bond formation so this is what it is actually going to happen. So initial the first codon is actually going to be present on the P site and then subsequent codon will come on to the A site and that is how it is actually going to be start doing the synthesis. So by sequential addition of the you know tRNA which actually contains the amino acids it is actually going to be participate into these peptide synthesis. So the enzyme peptidyltransferase catalyze the formation of the peptide bond the activity of this enzyme lies on to the 20S ribosomal RNA of 60S ribosomal subunits. It is therefore the rRNA not the protein referred to as ribosome that catalyze the peptide bond formation.

Next result of the peptide bond formation is the attachment of growing chain to the tRNA in the A site. So in this you know cartoon what you see here is that this is actually A site and this is actually the P site. So what you see here is that the codon is entering into the A site right it is sitting and then the peptide which is growing chain is actually been transferred on to this and then it moves on to the P site and from P site whatever is left over that will enter into the E site and it is actually going to exist. So this will continue as long as you have messenger RNA when it reaches to a place where you are actually going to have the stop codon then it will actually going to stop the synthesis of the peptide chains. So then you are going to have the translocation so ribosome move to the next codon of the messenger RNA this process involves the moment of the growing peptide chain from the A site to P site translocation require the elongation factor 2 and the GTP the GTP get hydrolyzed and supply energy to move on to the messenger RNA and the elongation factor 2 and GTP complex recycled for the translocation.

About 6 amino acid per second are incorporated in the course of elongation in the translation in the eukaryote. So this is actually the synthesis rate right where you have the 6 amino acid which are going to be synthesized in per second. So if I tell you the length of a protein they can actually be able to calculate how long it will take for a protein to synthesize remember that it does not include the transcription actually. So it is actually should be you know first the RNA is going to be transcribed and then after that it is actually going to be the this speed. So this is actually the diagram which actually shows how the elongation is going to happen in the eukaryotic system.

And then we have a termination so termination of eukaryote is almost similar with the prokaryotes which depends upon the eukaryotic release factors. So eukaryotic release factor or the RF 1 recognizes all three termination codons that is UA, UAG and UGA

and with the help of the RF 3 it terminate the translations. So upon termination the ribosome is disassembled and the complete polypeptide is released. The class 1 factor that is the RF 1 is responsible for the high fidelity stop codon recognition and the peptidyl-tRNA hydrolysis. The class 2 factor that is the RF 3 is translational GTPase that is more closely related to the EFTU than EFG and RF 3 accelerate the peptide release and increase terminal efficiency at stop codon in a manner that depends upon the GTP hydrolysis.

So this is what exactly happened and it termination is exactly happened in the same way as we have discussed about the prokaryotic system. So these are the some of the steps what happened in the termination steps in the eukaryotic system. So one of the stop codon or terminal signal that is the UA, UGA and terminate the growing polypeptide chain and when the ribosome encounter the stop codon there is a no tRNA available to bind to the A side of the ribosome instead a release factor binds to it. Once you have the RF 1 recognizes all three stop codon and RF 3 stimulates the termination events. Once the release factor binds the ribosomal unit rival to the unit fall apart releasing the large and the small subunit the tRNA carrying the polypeptide is also released being upon the polypeptide product and the ribosome recycle occur at the end only in the eukaryotes.

So this is the ribosome recycling and or will say ribosome dissociation and then reassociation on to the next messenger RNA. So after the release of polypeptide and the release factor the ribosome is still bound to messenger RNA and it is left with the 2 deacylated tRNA. To participate in the new round of polypeptide synthesis these messenger RNA and tRNA must be released and the ribosome must be dissociated into the small subunit and to the large subunit that is what we have already discussed how the ribosome is getting dissociated with the help of the initiate so many eukaryotic initiation factors. Collectively these events are termed as ribosomes recycling so what happen is that when you are going to start with a start codon it is going to start then you are going to enter into the elongation phase and when it enter into the stop codon after the stop codon the ribosome is getting dissociated so it is going to it is going to dissociate the small subunit and the large subunit and then these small and large subunit are actually going to assemble onto the new messenger RNA. And once they are get dissociated this messenger RNA is also free for starting the new cycle of protein synthesis and that is how it is good for conserving the energy in terms of not synthesizing the same messenger RNA again and again that you can be able to reuse the same messenger RNA again for the synthesis of this particular protein and on the other hand you do not have to synthesize the ribosome also.

You can actually disassemble and then reassemble that onto the new messenger RNA on or onto this particular messenger RNA whatever the case these ribosomes are free for

starting the new protein synthesis cycle. So this is all about the central dogma of molecular biology where we have discussed about the translation into the eukaryotic system we have discussed about the translation in the prokaryotic system and what we have also discussed that how the eukaryotic system is different from the prokaryotic system and you might have seen that the eukaryotic system is much more complex much required many more elongation factors many more initiation factors and so on. So with this brief discussion about the translation I would like to conclude our lecture here in our subsequent lecture we are going to discuss about the post translational modifications. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 07
Translation
Lecture-32 Post-Translational Modifications

Hello everyone, this is Dr. Vishal Trivedi from Department of Bioscience and Bioengineering, IIT Guwahati and what we were discussing we were discussing about the different aspects of the molecular biology in this particular MOOC course. So, so far what we have discussed we have discussed about the cell biology, we have discussed about the different types of biomolecules, we have discussed about the role of these biomolecules into the cellular metabolisms and other kinds of important cellular functions such as cell division, cellular apoptosis, cell growth and the autophagy. In previous couple of modules we were discussing about the central dogma of molecular biology. So, we have discussed about the what is the requirement of the central dogma of molecular biology, why it is so important for the maintenance of the life on earth and what are the different processes which are being part of the central dogma of molecular biology. So, what you can see here is that the central dogma of molecular biology is consist of three important processes.

One is the duplication of the genomic content that is being done by a process known as replication followed by the synthesis of the different types of RNA from the DNA and that process is being known as the transcription and then from the when the RNA is been synthesized it is going to be utilized by the ribosome and the protein synthesis machinery to synthesize the protein and this process is known as translation. So, in the current module we are discussing about the translations and so far what we have discussed we have discussed about the translation different types of steps in the translation in the prokaryotic system and as well as in the eukaryotic system and we have discussed about the genetic codes, how the people have discovered genetic code and what are the different types of genetic codes which are possible and then we have discussed about the initiation, elongation and terminations in the case of the prokaryotic as well as the eukaryotic system and what are the different significant differences between the translation in the prokaryote versus the translation in the eukaryotes. Now with this detailed discussion about the translation in the current lecture we are going to discuss about what happens once the protein is been synthesized and how the cellular system is changing this protein. So, that it becomes more useful for regulating the different types of events or how it is participating into the different types of processes.

So, what we have discussed we have discussed about the cell tone dogma and at the end

of the translations you are going to get the synthesis of the different types of proteins. Now once you have a protein right it is actually going to be present as a nascent protein or I will say it will not be a un mature proteins right. So, this protein has to be converted into a protein which is more of mature right. So, what is mean by the mature is that this protein may require the some kind of modifications which is not possible by the translational machinery to it. So, that it will be more useful and it will be more functional in terms of the doing the different types of task what is been given by the cellular system or what would be the requirement of the cellular system.

So, that is why I am writing this as protein as a immature it is not applicable to all the proteins it is possible that some of the protein which are been produced after translation could be in the final form and could be utilized as such. So, one of the classical examples what you have seen is the proteins are getting truncated right. So, one of the modification could be the truncation. Classical example is when you are converting a pro enzyme into active enzyme right. There are many examples such as pro-pepsinogen right sorry pepsinogen right.

This is a inactive form of the enzyme and when it is getting cleaved off it is forming a enzyme which is called as pepsin and that is the active enzyme. So, this is also happens after the synthesis of the pepsinogen. Now truncation is a very very important component then you also have the different types of the covalent or the non-covalent modifications mostly the covalent modifications such as the phosphorylation such as glycosylation such as ubiquitylation or acetylation. So, all these are these different types of changes right and all these changes are bringing a very significant difference into that particular protein. For example, the phosphorylation it is going to impart a negative charge right and that negative charge is very important right because it may actually change the different types of scenarios and different types of interactions.

Similarly, the glycosylation that is actually going to impart a hydrophobic patch right. So, if you have a you know glycosylation it is actually going to reduce the solubility of that particular protein and because it is going to be less soluble it is actually going to increase the stability of that particular protein and we will discuss that in detail how it happens actually. Similarly ubiquitylation that is also part of the protein degradation pathway so it is actually going to be a part of the protein degradation pathway so it is actually going to induce the protein degradation and acetylation is also going to induce the negative charge and it is actually going to add the or its actually going to add up the energy part right. So, it is going to add up the high energy bonds and that is how it is actually going to provide the energy into the system. Same is happening with the phosphorylation also that is also actually going to add up the energy into the system and that is how it is actually going to incorporate the high energy into the system.

Now let us first talk about the phosphorylation right. So, phosphorylation is one of the most important and the important modifications after the protein is been synthesized. So, post translational phosphorylations. Phosphorylation is an important post translational modification it is prevalent from the bacteria to higher eukaryotes sustainably as many as mainly two types. First it acts to functionally regulate the catalytic activity of protein by defining a rigid and permanent three dimensional structure.

Secondly temporarily phosphorylation proteins serve as an anchor for other protein substrate in signal transduction pathways. As such it act as a key player in the regulation of many cellular processes such as cell cycle, cell growth apoptosis and regulation of the cellular transduction. So, what it says is that if the protein is going to be get phosphorylated. So, it is going to be get phosphorylated with the help of the ADP and the protein is going to contain the going to take up the phosphate from the ADP. Then this protein is actually going to have or it is going to incorporate two different types of functionality functional you know functional things into the protein.

It is actually going to regulate its activity and you will see examples right. And the second part is it is actually going to add up into the interaction of the proteins during signal transduction. I am sure you might have seen the cellular signaling. So, you might have seen different types of signal transduction pathways. I think we have discussed about the apoptotic pathways intrinsic pathway and extrinsic pathway and how the cell signaling is occurring from the receptor pathway or the mitochondrial pathway and how the proteins are involved into this.

So, phosphorylation is also having the different similar kind of cascades and that is been maintained by the phosphate. So, some once you have the phosphorylation it is actually attract the other proteins and that is how you going to have a detailed signal transduction. So, for example, the insulin signaling and there are so many growth factor signaling there are so many signaling pathways which are very well decide and very well discussed. Similarly you have the regulation of enzymatic activity so, sometime the phosphorylation actually brings the higher activity sometime the phosphorylation means the lower activity. And you might have seen many of these examples when we were discussing about the glycolysis and Krebs cycle because you might have seen the many of the enzymes present in the glycolysis and Krebs cycle whose activities are been regulated similarly by the phosphorylations.

Now, let us see what is the mechanism of phosphorylations. So, in eukaryotic cell the phosphorylation is known to occur only at the side chain of the 3 amino acids serine, threonine and tyrosine ok. And what you see here is that there is a significant similarity

between the 3 amino acids and what is the similarity? Similarity is that they are all containing a hydroxyl group and which contains a lone pair of electrons. So, this is very very important feature of an of a amino acid because the amino acid should have a lone pair of electrons so, that it actually can participate into the phosphorylation reactions. And that is why you will see that in the eukaryotic system you have the serine, threonine and tyrosine.

But in the other system like for example, in the bacterial system you also have the aspartate which is getting phosphorylated you also have the histidine which is getting phosphorylated and so on. So, because the histidine is also having the nitrogen which is having the lone pair and the aspartate is having the this right aspartate is having this as a lone pair right. This is the terminal hydroxyl, a carboxylate group and you are actually going to have the lone pair on this. So, that is what it is actually going to do. So, in the histidine you are actually going to have the nitrogen, in the aspartate you are actually going to have the COOH group and this is actually going to participate into the phosphorylation reaction.

So, as I said in the lone pair of electron is going to be you know the requisition or going to be important for these kind of modifications. This is because these amino acids harbour a nucleophilic hydroxyl group. The terminal phosphate group on the universal phosphate donor adenosine ATP serves as the point of nucleophilic attack from the OH group which results in the transfer of phosphate group to the amino acid side chain. Magnesium ion act as a catalyst by chelating the gamma and beta phosphate group resulting into the lowering of the threshold for the phosphoryl transfer of the nucleophilic groups. So, that is what you are going to see right.

So, this is the example of the serine kinase. You can expect the similar kind of mechanisms even for the other modifications. So, what you see here is that this is the actually the protein where you have the serine and then serine has actually this lone pair of electrons. So, this lone pairs of electrons are actually going to act on to the on to this particular phosphate. So, what you see here is this is the ATP and in the ATP you have the three different types of phosphate.

You have the alpha phosphate, you have beta phosphate and you have gamma phosphate. So, it all the phosphorylation reactions actually works with the help of the gamma phosphate. So, in the gamma phosphate there will be a nucleophilic attack from the lone pair of electron what is present and as a result this particular group is actually going to get broken down from here and then this will go like this. So, it is actually going to get converted into ADP and this portion will actually go and attach here and that is how you are going to have this the phosphoserine and that phosphoserine is

actually going to be formed. Now, the similar kind of thing can happen even with the
theronin or histidine and aspartate.

So, the enzymes name would be different in this case it is going to be called as serine kinase. If it is a aspartate then it is going to be called as aspartate kinase, if it is a histidine then it is going to be called as histidine kinase and so on. And majority as I said you know histidine kinase and aspartate kinase are more prevalent in the case of the prokaryotic system, but in the eukaryotic system you only have the serine, threonine and tyrosine. Now, these conformational changes can affect the protein in two different ways. So, the phosphorylation once you have a phosphorylation it is actually going to neutralize the or it is actually going to bring the additional negative charge.

What you see here is that right now it is actually uncharged it does not have any any charge this is the uncharged amino acid, but this one is negatively charged. So, it is actually bringing additional charge and because of this additional charge wherever the serine is present it either will make up the new interactions. So, it is actually going to contribute into the new interactions because negative is going to attract by the positive right. So, it is going to have the some kind of solvage interactions or hydrogen bonding interactions or van der waal interactions with the neighboring residues or it may actually break the interactions. Both there are both possibilities right it either can form the new interactions or it actually can break the pre existing interactions.

So, if it is going to bring the new interaction it is actually going to make the structure more compact and that is how or it may actually bring the conformational changes, but if it is breaking the interaction pre existing interactions then it is actually going to make the interaction make the structure more loose and majority of these interaction majority of these modifications always occurs in active site right. So, active site of an enzyme or sometime into the regulatory site also, but mostly if it is happen in the active site the active site could be more compact or active site could be more relaxed and both of these events are actually going to participate or going to modulate the catalytic activity of the enzyme. So, sometime they may not doing anything with the catalytic activity they may be actually bringing the additional you know surface right. So, the negative charge is there and it is actually going to bring the additional charge additional charge and that additional charge could be a would serve as a interaction site or docking site for new proteins right. This is what we said in the previous slide right that sometime it actually provides the docking site for the new interaction and that is how it is actually going to participate into the single transactions because until this phosphorylation does not occur this is not going to be recognized by the next proteins and once the next protein will go and bind this it is actually going to be get phosphorylated and so on and that is why there will be a cascade of the phosphorylation from one enzyme to another enzyme and the

third enzyme and so on and that is why it is actually going to make a single reductions.

So, once this happens it is actually going to bring the conformational changes and that may affect the protein in two different ways. Phosphorylation causing the conformational changes in the phosphorylated protein this conformational changes stimulate the catalytic activity of proteins. So, any protein can be activated or inactivated by the phosphorylation. So, this is what exactly happens when you have an inactive enzyme and if it is getting phosphorylated with the help of a kinase and that utilizes the ATP then it is actually going to be phosphorylated and that phosphorylated enzyme could be more active. Sometime it happens that and it may be other way around that you have active enzyme and then it becomes inactive enzyme, but that depends on the type of conformational changes it is bringing into the system and what could be how it is going to impact the structure of that enzyme.

Once you have the phosphorylated enzyme then there will be another set of enzyme which is called as the phosphatase enzyme and that phosphatase enzyme is going to bring the enzyme back into the unphosphorylated form and that is how it is actually going to shuttle between the two state one is active enzyme another one is the inactive enzyme and so on. So, phosphorylated proteins employ the neighboring proteins which have structurally conserved domain that distinguish and bind to the phosphor motifs these domains are specific for the diverse amino acids. Protein phosphorylation is a reversible post translational modification which is carried out by the kinases which phosphorylate and phosphatases which dephosphorylate to the substrate. These two type of enzyme make possible the dynamic nature of the phosphorylated proteins. So, that the balance concentration of the kinase and phosphatase is very important for the cell and it is also important for the catalytic efficiency of a particular phosphorylation site.

So, these two enzymes the kinase and the phosphatase are actually going to be always be present in a pair. So, you are going to have a tyrosine kinase you also going to have a tyrosine phosphatase. So, that when the tyrosine kinase is going to make the changes into any protein there will be a tyrosine phosphatase that is actually going to reverse these changes. So, that you are not going to have the only you know downstream signaling you also going to have the upstream signaling. So, that you are also going to nullify the effects otherwise you can imagine that if you have started a process it has to be shut down right and that shutting down is only by the these kind of pair of these enzymes.

For example, if the kinase is making the system on then the phosphatase is actually making a system off and that is what it is actually going to work in pairs. Now, the question comes what could be the ways in which you can be able to detect the phosphorylations. So, you can actually be able to use the activity which is called as

western blotting. So, you can actually be able to use the western blotting and you can use the these modification specific antibodies like for example, you can use the anti phosphoserine antibody although we have not discussed western blotting so far, but I think we are going to discuss. So, that time you will be able to understand full detail about how the people are doing the western blotting and how the people are using these antibodies and that anyway we are going to take up when we are going to talk about the molecular techniques in the subsequent modules.

So, you are going to use the anti phosphoserine as the primary antibody right. So, many of these terms many of these things you will understand once we are going to discuss about this. So, the first is you are actually going to do it with the western blotting and you are going to use the specific primary antibodies such as anti phosphoserine antibodies anti threonine phospho threonine antibodies or anti phosphotyrosine. So, these antibodies are being directed against the phosphorylated tyrosine. So, it is going to recognize only the phosphorylated tyrosine this is not going to recognize the unmodified tyrosine present in the proteins and that is how it is actually going to give you the signal.

Apart from that you can also use the specific proteins like anti phospho p 38 or those kind of proteins also those kind of antibodies also. Then the second is you can be able to change you can be able to use the spectroscopic method right. So, you can use you can be able to determine whether there is a modification of hopsville and you can actually be able to use the many type of dyes which are specific for the phosphorylations. So, you can use like you can use the malachite green. So, malachite green is specific for the phosphoproteins it does not react with the normal protein and that is how you can actually be able to use malachite green.

Number 3 you can also use the fluoromit fluorescence or fluorescence spectroscopy. So, or you can use the so here you can use the a probe which is been called as pro q. So, you can use a probe which is called as pro q pro q is a proprietary probe which is available from the molecular probes and when it reacts with the phosphoproteins it when it reacts with the phosphoprotein not the normal protein it actually going to give you the orange fluorescence. So, what you can do is you can just resolve the proteins on to a SDS page. So, it is actually going to give you the bands right and these bands you can actually be able to stain with pro q and when you do that it is actually going to give you the pattern right it is going to give you the pattern wherever you have the phosphorylation.

So, where your band is phosphorylated it is actually going to give you the signal into the orange fluorescence and that can be captured with the help of the scanners. So, you are actually going to have a specific scanners which can actually be able to capture the fluorescence at the particular wavelength. So, it is going to give you the image right and

that can be used very easily and very with a lot of you know there is no it is a user friendly techniques of doing that. Then the fourth is you can also use the NMR you can use the other kinds of analytical techniques. So, that I am not going to discuss in detail because we have discussed that in a one of our MOOCs course which is called as experimental biotechnology.

So, we have discussed in detail about how the phosphorylation of the protein can be detected and so on. Now phosphorylation is playing a very crucial role in many of the events and as I said you know when you have a protein which is getting phosphorylated and you are actually generating a phosphoprotein you are actually changing many things right you are changing the charge right. So, you are actually adding up the charge then you are also changing the energy level right or you are actually adding the high energy bonds into the system you are actually changing the size of the protein you are also changing the molecular weight because it could be attachment point for the other protein and that is how they may actually come and attach. Size means it is going to change the hydrodynamic surface area of the protein because the protein could be more compact or less you know it is going to be loose and depending upon that it is going to change the size right and it also going to change the activity of the enzyme. So, it is going to change the many things and phosphorylation is a very very very important post translational modifications.

Now let us move on to the next phosphorylation modification and the next modification is called as the glycosylation. So, glycosylation is the dire function of the biosynthetic secretory pathway in the endoplasmic reticulum and the Golgi apparatus. Approximately 50 percent protein characteristically expressed in a cell go through this alteration which involves the covalent addition of the sugar moieties to the specific amino acid. Mostly soluble and membrane non protein expressed in the endoplasmic reticulum undergoes glycosylation including all secretory protein surface receptors and ligand. Moreover some proteins are being transferred from the Golgi to the ecosystem for their glycosylations.

So, glycosylation is a very very important post translational modification because it actually been used as a system to for the vesicular trafficking. So, it is actually going to be used as a system to deliver a protein from the one compartment of the cell to the another compartment of the cell. So, you know that the protein are actually going to be synthesized onto the ribosome and these ribosomes are actually going to be attached onto the endoplasmic reticulum and that is why you are going to have two different types of endoplasmic reticulum. You are going to have the smooth endoplasmic reticulum and you are going to have the rough endoplasmic reticulum. So, in the rough endoplasmic reticulum the ribosomes are attached onto the ribosome and attached onto the

endoplasmic reticulum and when they synthesize the protein, the protein goes inside the endoplasmic reticulum and get folded.

And once it get folded, it is going to be packed into the vesicle and then it is actually going to be transported to the Golgi bodies. When it goes into the Golgi bodies, the Golgi bodies are adding the specific glycosylation pattern which involves the glucose, fructose, mannose, arabinose and all other kinds of drug, all other kinds of glucose and a combination of glucose. And that is how it is actually going to make the specific signal. It is going to make the specific tag and that specific tag is actually going to be recognized only by the specific receptors, what are present onto a particular organelle. For example, if the Golgi bodies has put the tag for the mitochondria, then that tag is and then it is actually going to be packed into a vesicle and then vesicle will go and interact with the endometacontria and that vesicle is actually going to be taken up by the mitochondria by the interaction of a tag and the receptor what is present onto the mitochondria.

And then at the end this protein is going to be delivered specifically to the mitochondria, not to the any other compartment, not to the glycosines or not to the plasma membrane or not to like nucleus or any other thing. So, this is a very, very important glycosylation. So, glycosylation is actually a very, very important modifications. Now, the glycosylation is being done by the different types of sugar. So, it is actually can be glucose, it can be fructose, it can be few course, it could be mannose, galactose, arabinose and so on.

Now, once you add a sugar, you are actually going to add the hydrophobic groups. Remember that the sugar is polar, but the sugar moieties are when they coming together, they are not very polar. So, they are actually going to bring the dual characters, they are actually going to bring the hydrophobic charges also and they are also going to bring the polar charges also because if you see size structure of the sugar, this is the sugar and what you see here is that it is going to have the OH, it is going to have H, but also going to have this ring. So, this ring is actually hydrophobic and these are actually the polar groups. So, it is actually going to bring the mixed environment into the interaction and that is how it is actually going to make the proteins, there are it is going to make the modifications.

Then the second is it is actually going to change the size because these are actually not going to get compact. When they are being present inside the protein structure, they are actually going to make the protein structure more loose and because of that it is actually going to impact the size of the protein. Then it also going to change the interaction of the protein with other proteins, with proteins or with substrate because you know if there is a

glucose molecule which is going to be attached, it is actually going to occupy that space and it is not may not provide the space for the substrate to interact. Similarly, it may actually allow the interaction of the some of the proteins. And glycosylation is also going to change as I said you know the surface chemistry or the charges of what are the present in.

So, that is what now the question comes how you can be able to study the glycosylation if the protein is getting glycosylated. Number one you can actually be able to use the western blotting with the help of the anti glycosylated antibodies and that anyway we are not going to discuss. Number two you can actually be able to use the SDS page. Remember that this method what we are discussing is more about the if you have the pure proteins. So, suppose you want to know whether the protein is getting glycosylated or not.

So, when you what you can do is you can add the protein you can add the glucose. So, what will happen is it is actually going to make the glycosylated protein. Now, whether this is forming or not that you can actually be able to know because it is it could be one single glucose or it could be a multiple glucose. So, when the glucose is going to bind it is actually going to change the size of the protein it is not going to change the molecular weight to that extent. So, it is not going to change the molecular weight it is actually going

to change the size.

So, if the protein is very compact it may actually make the bigger balls. So, when you run it on the SDS page for example, so this is the protein and this is going to be the protein which is glycosylated or you have run the reaction for glycosylation. So, imagine that if you have a protein band here right then because the size is gone up right. So, it will behave like as a high molecular weight protein and that is how it is actually going to run

like here.

Now, how you know that this is actually glycosylated. So, what you can do is you can actually be able to treat this with acid or you can treat it with the enzyme. There are specific enzymes like the glycosidase. So, if you treat it with the glycosidase right then what will happen is this is actually going to be get converted into proteins. So, if I have another lane right if I have another lane of saying that protein glycosylated minus plus enzyme or acid then what will happen is this band is again going to return back to the normal band. So, this is actually the protein band this is the glycosylated band and this is the

unglycosylated band.

So, this will prove that the protein is glycosylated. If it is does not move if it is remain here then it is actually been a some kind of artifacts or it is not getting specifically been

glycosylated there are something happened funny and it cannot be explained by this. So, there are many other method you can actually have the some of the dyes periodic acid and all the so that anyway we are not going to discuss in detail about this. Now, the second another post translational modification is the ubiquitylation. So, this is another post translational modification where ubiquitin which is a protein actually is added to the protein. Ubiquitin is the eukaryotic protein coded by the 4 different genes in the mammalian cells such as UBA 52, RPS 27A and UBB and the UBC.

Protein is made up of 76 amino acid and has a molecular weight of 8.5 KDA. It is characterized by the specific presence of C terminal tail and the 7 lysine residues. In ubiquitylations basically the carboxylic acid side of the terminal glycine from the diglycine in the active activated ubiquitin forms an amide bond to the epsilon amino of the lysine in the modified protein. It marks the cellular protein for the process of degradation via the proteasome changes the protein localization prevent or promote the protein-protein interaction. So, what happened is that in the protein you have the different types of lysines.

So, when you are adding the ubiquitin and ubiquitin has a glycine. So, glycine on the C terminal and when you are going to do the ubiquitylation reactions what will happen is that the protein is having a lysine and it is going to attach by glycine and the ubiquitin is going to be attached. So, this ubiquitin is going to be attached with the help of the terminal lysine what is present on the ubiquitin by binding to the lysine and that is how it is protein is actually going to be tagged for the cellular degradation. Now, how it happens that you are going to you know ubiquitylate this protein not any other protein. So, there is a complete set of reaction what are going to happen when you are going to do ubiquitylation of a particular protein and there is a balance and checks which actually allows and to the identifier that particular protein and then only it is actually going to be get ubiquitylated.

So, it actually happens in the multiple steps. So, in the step one you are going to do the activation of ubiquitin. So, this is going to happen here. So, it occurs in a two step reaction process at first the ubiquitin interact with the ATP and forms a ubiquitinate adenylate intermediate. In the next step the ubiquitin is transferred to E1 active site enzyme containing the cysteine residue. This causes the formation of thioester linkage between the C terminal carboxyl group of the ubiquitin and the E1 cysteine cell filter groups.

So, this is what exactly happens. So, this is the step one where the this is the ubiquitin and it is actually going to be adenylate with ATP and that is how it is actually going to form the activated ubiquitin and that activated ubiquitin is going to bind the E1 ubiquitin

and it is actually going to you know form the thioester linkage. Then in the step two the step two there will be a transfer of ubiquitin E1 active site to the E2 active site via the transesterification reaction. So, that is what exactly happens then you are going to have the entry of the E2 ubiquitin and it is E2 enzyme and it is actually going to replace the E1 enzyme and it is going to form the E2 with the thioester linkage the ubiquitin is present. In the last step of the ubiquitin lylation cascade there is a formation of the isopeptide bond between the lysine of the target protein and the C terminal glycine of the ubiquitin via the activity of the one of the hundred of the E3 ubiquitin ligase. So, then you are going to have the E3 ubiquitin lysase and it is actually going to bind on one side E2 which actually has the ubiquitin and the target protein and the other side you are going to have the proteins which actually has the I mean lysine groups and then this is going to be transferred onto the amino group right of the lysine and that is why you are going to have the ubiquitin which is attached onto the target protein by the isopeptide bond.

So, this is actually a peptide bond but because you have a main chain peptide bond this will be called as the isopeptide bond okay. And you can have this cycle continue so from here it goes again here and then this cycle continues and then you are going to have the addition of multiple ubiquitin molecules. So, you can imagine that if you have a protein, protein may have the lysine and to this lysine the ubiquitin is going to be attached in this first cycle in the second cycle another ubiquitin is attached in the third cycle another ubiquitin is attached and so on. And remember that I told you right the ubiquitin also having the lysine. So, it is actually possible that sometime you may it may have the ubiquitin like this.

So, in some cases it happens that instead of adding the ubiquitin to the target protein it may actually add the ubiquitin to the existing ubiquitin because the ubiquitin has a 7 lysine residues the terminal C terminal amino acid. So, it is actually can attach 7 ubiquitin at once right and that is why the ubiquitin elation could be mono ubiquitin elations or it could be poly ubiquitin elations okay. And you can see that how it is actually going to change. So, you can have a mono ubiquitin elation. So, that only the one ubiquitin is going to attach onto the substrate protein or it could be a poly ubiquitin elations.

So, poly ubiquitin elation could be linear chain or it could be branch chains right. So, it could be branch change and the purpose of the ubiquitin elation is that it is actually going to mark the protein for the cellular degradation. So, it is going to deliver that protein into the proteasome systems. So, what happen is that when you are going to have a protein and when you are going to do the ubiquitin elations. So, if we say getting the ubiquitin elations it is going to be get converted into the protein ubiquitin okay.

So, I am just making taking a simple example of the mono ubiquitin elations okay. Now, ubiquitin is a 8.3 kDa protein okay. So, if the one protein is been attached it is actually going to change the molecular weight of the protein by 8.3 kDa okay. So, if it is a mono ubiquitin elation if it is a poly ubiquitin elation you can actually be able to you know multiply the same number and so when there will be a ubiquitin elation it is actually going to change the molecular weight of the target protein.

Number 2 it is actually going to change the size of the protein. Then number 3 it is actually going to change the surface interactions right. So, it is actually going to allow this protein to interact and because it is changing the molecular weight it is also going to impact the solubility of this particular protein because more and more protein you are adding it is becoming more and more heavy and if you know that heavy molecules are going to be less soluble compared to the light molecules right. Because the protein itself is soluble but when you are adding all these kind of thing it is going to be more soluble it is going to be less soluble. So, it is going to impact the solubility this means once it can impact the solubility it is actually going to make the particulate matter it is going to be form the particulate matter and these particulate matters are actually going to be a part of the proteasome degradation pathway. So, it is going to be attract by the proteasome system and that is how they will be delivered to the proteasome and they will be get degraded by the system.

So, it is going to be degraded. In some cases the monoubiquitylation or polyubiquitylation because it changes the size and the surface interactions it also brings the additional functionality into the system. Now how you can be able to study the ubiquitylation? So, you can actually be able to have the as I said you know you are actually going to have the change in the molecular weight, you are going to have the change in the hydrodynamic surface area which means the size and it is also going to change the surface chemistry. So, first thing is you can actually be able to do the western blotting with the help of the specific antibody that we are not going to discuss number 2. You can also be able to check the molecular weight. So, you can actually be able to run the SDS page and SDS page is going to tell you because in the each ubiquitylation round it is actually going to increase the size by 8.

3 kDa. So, you can imagine if you are going to have the addition of 5 ubiquitin then it is actually going to be approximately 40 kDa. There will be change in the 40 kDa and so on. So, if you run the system on the SDS page and suppose this is a protein and this is a protein ubiquitin, then it is actually going to change the pattern. So, it is going to change the pattern and remember that with the ubiquitylation because you are going to get multiple species, you are going to have suppose this is a protein.

So, we are going to have a ubiquitin then you are going to have this. So, this is in the cycle 1. And again in the same protein, you are going to have ubiquitin then it is going to attach with another ubiquitin right, it is going to be another ubiquitin this is the cycle 2. So, in the cycle 2 you are going to have 2 ubiquitin then cycle 3 you are going to have 3, 4, 5, 6 like that. So, that is why you are not going to get single band you are going to get a band here then band here band here band here band here. So, it is actually going to show you a ladder kind of pattern or I will say it is actually going to show you the sphere.

So, if you are getting a smear then that means that it is actually going to be ubiquitination. Now, how you are going to do the confirmatory experiment. So, what you can do is remember that the addition of the ubiquitin to the protein is been done by a bond which is called as the isopeptide bond right. See this isopeptide bond cannot be cleaved by the normal proteases ok, it is actually going to be cleaved by the cysteine proteases.

So, it is very sensitive for the cysteine proteases. So, what I can do is I can just take the PU and I will add the cysteine protease ok. So, if I add the cysteine protease this all ubiquitin is going to be removed and ultimately I am going to see a electrophoretic mobility of this protein to its original level. So, this is the protein, this is the protein plus ubiquitin and this is the ubiquitin we have lost the ubiquitin. So, all the ubiquitin you will see here actually you are going to see some of the ubiquitin smear on the bottom of the particular gel.

So, this is one of the method through which you can be able to detect the ubiquitin. Number 3 you can actually be able to do the gel filtration and as I said you know it is going to change the size of the protein. So, you can actually be able to use the gel filtration chromatography to know the size of the protein. We are not going to get into the detail of these because all these we have discussed in detail in my another MOOCs course which is called as the experimental viotechnology. So, if you are interested to know more about this you can be able to go through with this particular MOOCs course and it is actually going to tell you the complete detail about how the chromatography technique can be used for studying the ubiquitin dilutions.

Now, let us move on to the next modification and the next modification is the methylation. This process refers to the addition of methyl group to the nitrogen or the oxygen to the amino acid side chain and methylation is irreversible whereas, o methylation is potentially reversible. So, methylation changes hydrophobicity of the amino acid and utilizes the negative charge when attached to the carboxylic acid. Many

methyl main methyl group contribution for such reaction is SAM or acetetyl methylene. This reaction is mediated by the enzyme which is called as methyl transferase and methylation process is involved in a epigenetic regulation as the histine methylation and demethylation. So, this is the methylation we have discussed in detail when we were talking about the chromatin romaned and genomic DNA and how the packaging of the genomic DNA occurs into the eukaryotic cell.

So, this is all about the post translational modifications and what we have discussed so far we have discussed about the central dogma of molecular biology where we have discussed about the replications transcription and translations. We have discussed we have tried to discuss in detail of this different processes into the prokaryotic as well as the eukaryotic system.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 08
Molecular Techniques (Part 1)
Lecture-33 Southern Blotting

Hello everyone, this is Dr. Vishal Tewedi from Department of Biosciences and Bioengineering IIT, Guwahati and what we were discussing we were discussing about the different aspects of the molecules in the course molecular biology. So, so far what we have discussed we have discussed about the cellular structures in the initial modules then followed by we have discussed about the different types of biomolecules. So, we have discussed about the DNA, RNA, proteins and then we also discussed about the enzymes and following that we have also discussed about the different types of cellular processes. So, we have discussed about the central dogma of molecular biology and then we have seen the detailed discussion about the different types of processes what is occurring within the central dogma of molecular biology.

So, we have discussed about the replications, transcription and translations. And with this kind of discussion we understood that the molecular biology has developed the different types of basic principles on which you can be able to develop the different types of techniques and these techniques can be used even to explore the aspects of the molecular biology or they can also be used for different types of applications. So, in today's lecture we are going to discuss about the some of the important molecular techniques and how these molecular techniques can be used for answering some of the basic questions related to the cellular physiology or other aspects related to the molecular biology. So, when we talk about the molecular biology, the molecular biology is all about the molecules right and these molecules interact due to a specific requirements.

I am sure you can recall when we were discussing about the proteins. We said that the protein is interacting with the substrates and protein is interacting with the other proteins due to the recognition of the complementary regions due to the specific three dimensional requirements, geometric constraints, stereo specificity and other kinds of requirements. And because of these kind of specific requirements the molecular biology or the interaction between the molecules can lead to the development of different types of techniques. So, when we talk about molecular technique, molecular technique can be of many types ok. Molecular technique where you are actually going to amplify a particular DNA or other kinds of regions.

So, it can be a amplification technique. Within this amplification technique you can

have the different types of amplification techniques such as the polymerase chain reactions, you can have the real time PCR or you can also have the rd PCR ok. So, this is I will write real time PCR and this is the rd PCR or reverse transcriptase PCR. And all of these amplification techniques are having the utility in terms of the amplifying a specific region of the genome. So, that you can be able to get the genome fragment and these fragments can be used for multiple for the downstream cloning applications and all that.

So, that anyway we are going to discuss in detail when we are going to take up the molecular cloning. Then apart from that you can also have the different types of the blotting techniques. So, within the blotting techniques you can have the techniques which is called as southern blotting. You can have the northern blotting and you can also have the western blotting and you can also have the eastern blotting ok. So, most of these blotting techniques where you are actually going to use the hybridization techniques and you are actually going to blot the molecule of your interest onto a support system and then you are going to detect that with the help of the different types of the analytical reagents.

For example, in the case of southern blotting you are going to use as the DNA as a probe and northern blotting also you are going to use the DNA as a probe and then western blotting you are going to use the different types of antibodies and so on. Apart from that you can also have the molecular techniques like where you are actually going to have the sequencing technique. So, you can actually be able to do the sequencing of the DNA and protein also. So, you can have the sequencing of the DNA or you can also have the sequencing of the proteins and that can be useful in terms of the identification of the that particular region or accuracy of the cloning and other kinds of things. And there are many techniques for example, you can have the RFLP, you can have the RFLP you know you can have the small molecular polymorphisms and all those kind of things.

So, the so molecular technique is a very vast subject and there are there are there could be a MOOCs course on that actually. So, what we are going to do is within this molecular technique topic what we are discussing we are going to discuss about the amplification techniques because they also going to have the downstream applications into the molecular cloning and then we are also going to discuss about the blotting techniques. So, in current module we are discussing about we are going to discuss about the blotting techniques. So, within the blotting technique we are going to discuss about Northern blotting, Northern blotting and Western blotting. And in the next module we are going to discuss about the amplification techniques such as the PCR, RT-PCR and real-time PCR.

And all of these techniques will help you to understand the application part that will also we are going to discuss at the end of this particular course. So, as far as the blotting technique is concerned the blotting techniques are been categorized based on the target molecules. So, for example, as I said to know that we have the four different types of blotting technique we can have the Southern blotting, we can have the Northern blotting, we can have the Western blotting and we can also have the Eastern blotting. And the Southern blotting you are going to use the DNA as the target molecule. So, you are going to identify the DNA if DNA is a molecule then it is going to be called as the Southern blotting.

If you are going to detect the RNA then it is going to be called as the Northern blotting and if it is the protein then you are it is going to be called as the Western blotting and if it is a carbohydrate then it is going to be called as the Eastern blotting. Although Eastern blotting is not very popular, but since they were having the Southern, Northern and Western blotting the people have whatever the technique is using for detecting the carbohydrate onto the nitrocellulose membrane or other kinds of support system that is been called as the Eastern blotting. So, Eastern blotting is not very popular and that is why we are not discussing about Eastern blotting in this particular course. So, we will start with the Southern blotting then we are going to discuss about the Northern blotting, then we are going to discuss about the Western blotting and then lastly and so when we discuss about these techniques we are going to talk about the basic principle of the technique and then we are also going to talk about the underlining mechanisms and we are going to talk about how to perform this particular technique. So, let us start with the Southern blotting.

So, Southern blotting is a six step blotting technique and remember that in all the blotting techniques you are going to blot the material onto a support system such as the nitrocellulose membrane or the PVDF membranes and then you are going to perform the reactions on that particular support system. So, the first step is for example, so Northern blotting is for DNA and the purpose of Northern blotting is to detect a fragment of DNA into a particular genome or a genomic library. So, the first step is that you are going to have the genomic DNA of that particular organisms and you are going to isolate the genomic DNA from that organism. Then in the second step you are actually going to do the digest the genome because genome is very long. So, you cannot analyze that where the actually the probe is binding.

So, what you can do is you can digest the genome so that you are going to have the different types of fragments. And then in the third step you are actually going to use you are going to separate these DNA fragments onto a suitable agarose gel. So, you are going to use that sub suitable agarose percentage of the agarose gel and then you are

actually going to analyze that and that is actually going to give you the separate fragments. Then in the step 4 these DNA fragments are going to be transferred onto the nitrocellulose membrane with the help of the different types of transfer techniques. You can use the manual method or you can use the high vacuum method as well.

Once the it is been done then you are actually going to prepare a probe and you are going to prepare a radio labeled probe. Probe is a small fragment of DNA that you are actually going to use to detect this specific DNA into the genomic DNA. So, you are going to use the hybridization with the radio labeled probe and then in the step 7 you are actually going to do a washing and then you are going to do the autoradiogram. And at the end of the autoradiogram you are actually going to get the result and that result can be interpreted in terms of how many different DNA fragment are present into the genome and how many different copies. So, basically the with the help of the southern blotting you can be able to say whether this particular DNA fragment is present in that particular genome or not.

And the second information what you can also extract is what will be the copy number how many how many fragments or how many different gene of that particular fragment is present in that particular genome. So, it is actually going to tell you the location of that DNA fragment and it also going to tell you the information about the copy number. So, we have so far we have not discussed about the copy number, but I think we will going to discuss when there will be a molecular cloning and other kinds of events. So, the first step is that you are actually going to start with the genomic DNA isolations. So, genomic DNA isolation is also a step 1 in the genomic DNA isolation.

So, genomic DNA isolation is also a multi step process in that multi step process the first step is that you are going to do the lysis of the cell with the help of the detergent containing the lysis buffer. Now at this stage you can have the two different types of sources you can have the either the cell or you can actually have the tissue. For example, let us take an example of the liver. So, for example, if you have a liver from where you want to extract the genomic DNA and then you want to ask the question how many copies of that particular gene is present into the liver genome cell we can take an example of the cell also. So, for example, hepatocyte.

So, whether you use the tissue or whether you use the cell the ultimately is that if you are trying with the tissue then it has to be fragmented and into such a case that it should reach to a cellular level. So, tissue is made up of the different types of cells. So, what you are going to do is you are going to use the enzymes which actually can dismantle the inter space inter cellular spaces. For example, you can use an enzyme which is called as collagenase. So, if you treat the liver with collagenase collagenase is an enzyme which

actually degrades the collagen fibers and you know that the most of the cells are stick to each other because they have a matrix in between and this matrix is made up of the collagen or the collagen like material.

So, in the liver you have the collagen like material which is actually sticking these cells. So, if you treat them with the collagenase what you are going to get is you are going to get the cellular suspension or cell single cell suspension and this single cell suspension can be used for further downstream. So, single suspension means from the liver you got the single suspension and now you can actually be able to treat them with the help of the detergent. So, whether you are using with the single cell or whether you are using the tissue at the end you are going to get the cells and then you are actually going to treat them with the detergent and once you do with the detergent treatment the cell is actually going to break open then it is actually going to release the cellular content. And, what is there in the cellular content? The cellular content is actually going to contain the cellular proteins, it is going to have the RNA's, it is actually going to have the lipids and it also going to have the proteins and it also going to have the genomic DNA.

So, now from the cellular content your job is that you want to isolate the DNA, but you want to get rid of you do not want that. So, you want to get rid of the protein, you want to get rid of the lipids, you want to get rid of the other materials right. So, that is what you are going to achieve in the subsequent purification steps. Now, in the step 2 you are actually going to incubate the cells with the digestion buffer containing the protein SK, SDS to release the genomic DNA from the DNA protein complexes. Remember when we were discussing about the genetic material we said that in the eukaryotic system the genetic material is associated with the protein in the form of the nucleosome and these nucleosomes are even forming the higher order organizations right such as the chromosomes and other kinds of things.

So, the first step is that you want to isolate or you want to dislodge the protein part. So, that the DNA would be free and then you can be able to get the purification of the DNA from the cellular material. So, that for that what you are going to do is you are going to incubate the cellular content or the live cells with the digestion buffer which contains the protein SK and SDS. So, in the presence of SDS the protein is actually going to get denatured right and because the protein is going to denature it is actually going to lose its affinity for the DNA. On the other hand the protein SK is a protease right.

So, it is actually going to degrade the protein into the peptide or peptide fragments right. So, ultimately what you are going to get is you are going to get and peptide fragment can further be digested. So, if you continue the digestion it is actually going to get converted into amino acids right because the protein SK is a non

specific protease. So, it is actually going to start chewing up the proteins and ultimately what you are going to get is you are going to get the DNA which is separated from the DNA protein complexes plus you are going to have the protein which you are going to get from the cellular content and then you are also going to have the lipids right because lipid is also present in the cellular material and some minor component of like for example, you can also have the RNA right. So, now what you have is you have a mixture of DNA protein lipid and RNA and your target or target molecule is the DNA.

So, you can actually devise a strategy so that you can be able to purify the DNA, but get rid of the protein lipid and RNA species. So, that you can do sequentially by removing the protein lipid and RNA and that is how you are going to get at the end you are going to get the purified DNA. So, in the step 3 you are actually going to isolate the genomic DNA by the absolute alcohol precipitation. So, therefore, what you are going to do is you are going to add the absolute alcohol to the thing and it is actually going to precipitate the genomic DNA, but when it is going to precipitate the genomic DNA is it is actually so it is going to give you the genomic DNA, but it also going to have the small amount of proteins it is also going to have the small amount of lipids and it also going to have the small amount of RNA specially the messenger RNA. So, these has to be get rid of right this has to be get rid of.

So, in the next step what you are going to do is you are going to do the purification of the genomic DNA with the help of the phenol chloroform isomer mixture. And when you mix the phenol chloroform one mixture has two phases right you are going to have the aqueous phase and you are going to have the organic phase. See here you have the two phases you have the chloroform phase which is going to be the part of the organic phase right. And you know that when we are going to do a fractionation with the solvents the molecules will actually go to their respective phase right. For example, if there is a organic material then it actually going to prefer to the organic phase and if it is a polar molecules then it will go to the polar environment right.

And as a result what will happen is that in a append off you are going to have the two phases you are going to have the organic phase and you are going to have the aqueous phase. So, this is going to be aqueous phase and this is going to be organic phase. So, all the molecule which are soluble into the organic solvents such as lipids such as the other kinds of molecules they will be present in this whereas, the DNA and RNA is going to be present into the aqueous phase. So, in this step the phenol is actually going to denatured the remaining proteins and keep the protein into the organic phase. So, protein organic phase is actually going to take care of the lipids.

So, it is actually going to remove the lipids it is going to remove the protein because

protein is going to be denatured by the phenol and as a result the protein is also going to be present into the organic phase. Whereas in the aqueous phase you are going to have the DNA and you are also going to have the RNA because RNA is also going to be polar in nature. Now, next step is that you want the DNA, but you do not want the RNA. So, then you are going to do a further purification. So, first you are going to do a isolation of the genomic DNA what is present in the aqueous phase and it will again precipitated with the absolute alcohol.

So, if you do the reprecipitations it is actually going to precipitate the DNA and when it is also going to have the small amount of RNA and that you can actually be get rid of by putting the things into the RNA. So, if you incubate this mixture with the RNAs then it RNAs is actually going to chew up the RNA and it is actually going to destroy the RNA and then ultimately your genomic DNA is going to be isolated. So, this is what exactly is being depicted in this particular picture. First step you are going to do the lysis, second step you are going to do the enzymatic digestion, the third step you are going to precipitate the genomic DNA with the help of the absolute alcohol, but when you are going to precipitate along with the genomic DNA you are also going to get the lipids and the protein part and that you are supposed to get rid of with the help of a purification step. So, in that purification step you are going to add the phenol chloroform mixture and when you add the phenol chloroform mixture the phenol is actually going to denature the proteins whereas, the chloroform is actually going to remove the lipid part.

And when the phenol is going to denature the protein they also will be come into the organic phase. So, when you collect the aqueous phase this aqueous phase is actually going to contain the DNA and as well as RNA and then from this aqueous phase you can precipitate the alcohol with the help of absolute alcohol you can again precipitate the DNA and then you can treat this DNA with the small amount of RNAs. So, that you can get rid of the RNA part and then ultimately you are going to get the pure mammalian genomic DNA. Now when you run the mammalian genomic DNA onto agarose gel what you are going to see is that you are going to see a intact band which is going to be of high molecular weight. So, it is going to be of high molecular weight rather than a low molecular weight and this is going to be indication that you have isolated the genomic DNA.

If it is not a genomic DNA or if it is not an intact genomic DNA then in that case you are going to see a band, but below to that band you are going to see a streak or you are going to see a smear. That smear is actually an indication that you have broken the genomic DNA and you are actually going to generate the or you have degraded the genomic DNA. So, that is not good for the further analysis of into the sudden blotting. Now the next step is next step is step 2 you are going to have the generation of the

suitable size fragments because genomic DNA is very long. For example, when we were talking about genetic material in the eukaryotes we said that in the humans the length of the genomic DNA is approximately around 1 meter right.

So, if this kind of lengthy genomic DNA cannot be analyzed for you know for probing purposes. What you want is that you do digest that genomic DNA so that you are going to have the small small small fragment and then depending upon the binding of the probe it will actually going to say that if the probe is binding to the multiple location that means that many number of copies are of that particular fragment is present into the genomic DNA. So, for that purpose the genomic DNA has to be fragmented with a suitable system. So that you going to have not very big fragments not very small fragment, but should be a suitable size fragments and that you can be able to achieve by the two different methods. First you are going to do a restriction digestion.

So, genomic DNA can be digested with the frequent DNA cutting enzymes such as EcoR1, BamH1 or SAW3A to generate the random sizes of their DNA fragments. The criteria to choose a restriction enzyme or pair of restriction enzyme in such a way so that the reasonable size DNA fragment will be generated. As fragments are randomly generated and are relatively big enough it is likely that each and every genomic sequence is represented into the pool. So, the purpose is that you should not have very small fragments you should not have very big fragment it should be a relatively reasonable size. So that each and every genomic sequence should be presented which means it should be of such a size that a single gene can be present into that each fragment.

And because of that each fragment is actually going to represent the individual genes ok. And because of that if the you are going to see that the probe is binding at multiple location that means, your gene is having the that many copies present within the genome. Then the second method is the mechanical sharing. So, mechanical sharing genomic DNA can be fragmented using the mechanical sharing. Mechanical sharing method you can use the vertex you can use the sonication and so on.

But all these mechanical methods are need to be optimized very nicely so that you should not have very large fragments. Because large fragment had no use because it is going to have the representation of the multiple genes right. So, because of that it may actually going to give you the misguide. Now in this how you are going to perform the restriction digestion right. So, in the step 2 restriction digestion can be done with the help of the restriction enzyme.

So, what you need to do is you digest the 10 to 20 microgram of genomic DNA with the EcoR1 overnight in the presence of appropriate buffer and BSA. So, when you are going

to use any enzyme you require a suitable buffer and the other kinds of additives such as BSA and magnesium and all of the kind of things. Then take out the small aliquot from restriction digestion reaction and check it on to the 0.8 percent agarose gel. The presence of the smear with a small band visible band indicate the complete digestion of the genomic DNA and the suitability of the sample for the sudden blotting.

This means if I am going to run the genomic ah the restriction digestion in the control sample where I have not added a restriction enzyme I am going to see a intact band of the genomic DNA. Just now what I have shown you into the agarose gel. But if the cut fragment or restriction enzyme wherever restriction I am I have I am going to see a fragment and then I am going to see a complete smear kind of situation which means I am going to see a white band with the discrete band also. That means this is actually been done correctly and all these fragments are suitable for the sudden blotting and the step 2 is already been done. Now you can move on to the step 3 that means you are actually going to separate these DNA on to a suitable ah gel system.

So, the step 3 is the separation of DNA on to the agarose gel. So, separation of the DNA on to the agarose gel. Agarose gel electrophoresis is a standard gel to resolve the DNA. As a standard practice the agarose used for the sudden blotting is of ultra pure quality to avoid the contaminants affecting the migration of DNA. Observe the gel in the trans-culiminator and record the pattern in a gel documentation system.

So, this is very easy to say, but it is not easy. You are supposed to prepare an agarose gel then you are supposed to prepare the sample and then you are supposed to run it on to the agarose gel and then only you are going to see the pattern. So, for demo purposes I would like to take you to my lab where the students are actually going to show you the how to perform the agarose gel electrophoresis and how you can be able to observe the gel under the trans-culiminator and how you can be able to do the documentation. We have to analyze the results for amplification. For that we need agarose and TAE buffer.

First we have to weigh agarose and mix with the TAE buffer. It will not dissolve easily. So, we have to heat it in microwave oven until it get dissolved. Now agarose got dissolved in TAE buffer. We have to let it cool down up to 50 degree Celsius. Now before pouring we have to add ethereum bromide for detection purpose.

Now the gel got solidified. We have to take out the gel and keep it in the electroporotic apparatus. We have to gently remove the comb. Loose the knobs and keep the gel in the apparatus. Make sure that the buffer is submerged in the gel.

We have to fill the remaining part with 1X TAE buffer. Generally for analyzing the

DNA samples we will use agarose gel electrophoresis. This is the power pack and this is the electroporotic apparatus. This is the negative electrode and this is the positive electrode. We can change the voltage from here. For loading of sampling we have to mix TCR reaction mixture with 5X loading time. This is the TCR reaction mixture.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 08
Molecular Techniques (Part 1)
Lecture-34 Northern Blotting

Hello everyone. This is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering, IIT, Guwahati. And what we were discussing, we were discussing about the different aspects of the molecular biology in this particular MOOC source. And so far what we have discussed, we have discussed about the cellular structures, we have discussed about the basic structure and functions of the different types of biomolecules.

So we have discussed about the DNA, RNA, we have discussed about the proteins and then lastly we have also very briefly we have discussed about the enzymes. Following that we have also discussed about the central dogma of molecular biology and within the central dogma of molecular biology we have discussed about the replication, transcription and translation. And while we were discussing about all these, we have also discussed about the different variations between the prokaryotic and the eukaryotic system and other kinds of discussions. In the current module, we are discussing about the technical aspects of the molecular biology and how the basic principle of the molecular biology could have allowed to develop the different types of techniques and how these techniques can be exploited to answer the complex biological questions and as well as how they can be utilized for different types of applications.

So in the previous lecture if you recall we have discussed about the southern blotting and southern blotting is being used for detecting the DNA, a particular DNA into the genome. So in that discussion we said that it is a multi-step process where first you have to isolate the genomic DNA followed by the digestion of the genomic DNA so that you are going to have the different fragments and then these fragments are supposed to be separated onto a agarose gel followed by the transfer of these fragments onto the nitrocellulose membrane and then probing with the particular gene fragment which you are interested to identify and then ultimately you are going to develop the blot with the help of the autoradiogram. And the answer what you are going to get at the end of the southern blotting is that it is going to tell you that how many times this particular gene fragment or this particular fragment is been repeated within the genomic DNA and what would be the location of that particular fragment. So now in today's lecture we are going to discuss about the another blotting technique and that is called as the northern blotting. Northern blotting is been used for detecting the particular RNA into the RNA pool.

Mostly it is been done for the messenger RNA so you would like to answer which messenger RNA or which particular type of messenger RNA is been present. Northern blotting can be done to understand the expression of a particular gene. Does not give you the information about the translation but it is going to give you the information about the transcription. So it is understood that once you have the messenger RNA transcribed from the DNA it is going to be translated. So in this particular technique the basic steps remains the same that you are going to isolate the RNA and then you are going to separate the RNA onto the agarose gel and then you are going to transfer that onto the nitrocellulose membrane and then you are going to probe it with the suitable radioactive probe followed by you are going to develop that with the help of the autoradiogram.

So in a northern blotting you are going to have the six steps. So in the step one you are going to isolate the messenger RNA from the so you are going to isolate the total messenger RNA pool from the cell whether it is a prokaryotic cell or the eukaryotic cell. And then once you have isolated the messenger RNA then it is all these messenger RNAs are going to be isolated and run onto the agarose onto agarose. So you are going to separate the messenger RNA onto agarose gel so that you are going to get a pattern of the messenger RNA and then these patterns are going to be transferred onto a solid support so that you can be able to process because you know that the agarose is very fragile so you cannot process the messenger RNA fragment what are present onto the agarose gel. Instead you are going to transfer that onto a solid support such as the nitrocellulose membrane and then it is actually going to be easy for you to do it.

And once you transfer that onto the nitrocellulose membrane after that you are going to use the cDNA probe radioactive cDNA probe to probe the particular type of messenger RNA so cDNA will actually go and bind to its complementary sequence or complementary sequence or I will say complementary messenger RNA and that is how you are going to use this and then you are going to do the hybridization and you are going to do the development with the radio available probe and you are going to do a washing and the autoradiogram. So in the last step you are going to do the washing and the autoradiogram and ultimately what you are going to see is you are going to see the messenger RNA bands onto the onto the x-ray film and the intensity of these bands are actually going to give you the expression level or the transcription of that particular messenger RNA under the different experimental conditions. So what you can see here is that in this particular events you are actually going to have the similar kind of steps what we have discussed in the Southern blotting except that there you are actually the genomic DNA and then you are processing the genomic DNA with the help of the restriction enzymes and so on. Here you do not need to do that because your messenger RNAs are actually going to be individually separated right. So you just have to isolate the pool and then you are going to separate them onto the onto the agarose gel and then

you are going to transfer and all that.

So if you want to perform the Southern blotting you are supposed to have required the following materials. So you require the 1X SSC buffers. So the 1X SSC buffer is having the 5 millimolar sodium citrate pH 8 and the 150 millimolar NaCl. Then you also require the 5X buffers. So 5X XR is actually going to have the 0.

25 normal HCl, Tres-HCl pH 7.5, 0.5% sodium pyrophosphate 1% PVP, 1% BSA, 1% FICOL and 5% DSA. Then you also require the pre hybridization buffer and that will contain the formamide that is 2.5 ml, 5X buffer 1 ml, water 4.

4 ml and the NaCl as 0.292. And then you also require the salmon sperm DNA that is the blocking agent what you are going to use. So you are going to use that for. So this is actually going to be a part of the blocking agent.

Then you also require the Whatman 3mm filters paper and then you also require the labeled RNA or the DNA probes and you also require the UV lamps and the shaking water bath. Apart from that you also require the autoradiogram cassettes and you also require the autoradiogram x-ray films. Now before we discuss about the Southern blotting it is important to recapitulate and the re-review about the RNA structures. So if you see the RNA structure I think we have already discussed that because this is very important to understand then only you will be able to understand the messenger RNA isolations. So what you see here is that it is actually having the three important part.

One is you are going to have the 5 prime cap, the second is you are going to have the 3 prime UTR and the third is you are going to have the coding sequence. So this is the actually the messenger RNA what you are going to get and that is also responsible for the synthesis of the protein. But apart from this you are going to have the 5 prime cap and the 3 prime polyadenylation tail. So this polyadenylation tail is very very handy tool to isolate the total messenger RNA pool from a cell. So messenger RNA has the N-terminal cap structure, the coding sequence and the C-terminal polyA tail.

The nucleotide A form the two hydrogen bonding with the nucleotide T and this pairing is very specific. Exploiting this feature messenger RNA population can be isolated from the RNA pool using the polyT affinity chromatography. Remember that when we were talking about the RNA structures when we discuss about the different types of biomolecules, we discuss about the different types of RNA isolation methods. So we discuss about the trizole method and so on. Apart from that you can also use the affinity chromatography where you can actually be able to exploit this particular part of the messenger RNA and you can actually be able to utilize the fact that the A is actually

going to make the very specific interaction with the T.

So if you have a particular bead which actually contains the you know the T on that right. So if you have a bead which actually contains the T nucleotides then what will happen is that the A will actually go and bind to this and that is how the messenger RNA is going to bind specifically to these beads. Other molecules will not interact with these beads and that is how they can be washed away. So this method is called as the affinity purification of the messenger RNA. So affinity purification of messenger RNA and it can be done with the polyt-agarose beads.

So in a polyt-agarose bead what you are going to do is you are going to have the you are going to have the different steps. So in the first step what you are going to do is you are going to lyse the cells with the help of the lysis buffer. If it is a tissue then you are supposed to homogenize so that you are going to get the single cell suspension and after that you are going to lyse the cells and so that you are going to have the cell lysate. And this cell lysate can be put it into the binding buffer along with the oligo DT beads. So these oligo DT beads are actually going to have a bead on which you are going to have a linker and on this linker you are going to have the T nucleotide attached right.

So you have multiple T nucleotides what is attached to this and this T nucleotides are actually going to bind the A nucleotide very specifically because the A is having the specific base pairing to the T. So this T will interact with the A nucleotides and so it will actually going to capture all the messenger RNA which are actually having the A's right. So that is how you are actually going to have this is actually the messenger RNA and this is going to be the poly poly A tail right and this poly A tail is actually going to interact with the T or the T nucleotide what is present on to the beads. After that you are going to do the wash with the washing buffer and once you wash you are going to remove everything except the these bound messenger RNA. So you can actually be able to discard the supernatant and then you are going to do the elution so you are going to do the elution with the help of the competition and then you can respond this suspend the beads into the elution buffer and you can collect the separate the beads from the eluent with the centrifugation.

So when you separate you are going to have the messenger RNA into the aqueous phase and the beads are going to be get separated and then you can actually be able to get the pure poly A RNA molecule. So you are going to have the RNA species which actually contain the poly A tails actually. So in the steps what you are going to do is release the total RNA by lysis of our containing detergent or by the homogenization in the case of heart tissue then you mix the poly T containing beads with the total RNA species due to the mutual exclusive affinity the messenger RNA bind to the poly T then wash the beads

with the washing buffer to remove the non specific contaminating species and then you elute the messenger RNA from the bead and its purity can be checked on to the poly acrylamide gel. So in the step one you have isolated the messenger RNA pool from the cell after the treatment so you can have the multiple treatment you can have the untreated sample you can have a treated sample and so on. So you can actually be able to collect the messenger RNA from the these species.

Now the second step is that you are going to do the separation of messenger RNA on to the agarose gel. Now there you have a little complications because messenger RNA is a single standard nucleic acid and these single standard nucleotides are actually having the freedom to move around because if you have a double standard DNA like this then you have a restriction because all these are actually been bound to each other and because of that you cannot have you have lot of restrictions. So because of this you can actually may not be get the very good separation of these messenger RNA molecule on to the regular agarose gel instead you are supposed to run the denaturating agarose gels. Now why there is a need to run the denaturating agarose gel because the RNA gels are performed under the denaturating conditions and why it is so because the presence of secondary structure in the RNA allows the faster migration of RNA on to the agarose gel. So what happen is that if you have a single strand it can actually rotate and it can actually you know form the different types of secondary structures and because of these secondary structure it actually becomes very compact structure and once it becomes compact it runs very fast.

So if it is runs very fast it does not allow the agarose molecules you know to exert any kind of retardation forces and because of which the messenger RNA will not get separated instead of getting the individual bands what will happen is that you are going to get a smear because then you are going to get these no separations and that happens because the messenger RNA is going to form the different types of secondary structures they are going to form the stem they are going to form the hairpin loops they are going to form the pseudo knots bulge and they are also going to have the internal loops and multiple loops. The idea is that once you have these kind of secondary structures they are actually going to make the structure very compact and because of that they will run very fast. So as a result of these secondary structure it gives less time for the molecule to interact with the agarose gel and consequently less resolution within the different RNA species. Destruction of the secondary structure in the RNA structure minimizes these efforts and allow the better separation on to the agarose gel. So suppose you can imagine that you have RNA which is forming the secondary structure like this so if I and it is happening because you have the A here and you have a T here so you are actually going to have the you know the hydrogen bonding formed and because of that it may perform the different types of secondary structures like stems hairpin loop pseudo knots

bulge internal loops multiple loops and all that.

So if I break these bonds like if you add some kind of the for example if you add the strong salt so if you add the strong salt you are actually going to break the hydrogen bonding and because of that it becomes a straight line or it will actually going to adopt extended confirmation. Now if you are actually going to maintain the extended confirmation then they are actually going to follow the electrophoretic mobility principle and they are actually going to get separated based on the mass instead of the hydrodynamic surface area. So they are actually going to be get separated from the mass and then what happen is that when you run these you are actually going to get the separate masses and separate bands on to the agarose gel and that is what is required because in the subsequent step you are going to do the transfer you are going to do the hybridization and that is why it is important that you should know that where my probe is interacting whether it is interacting with this particular band or whether it is interacting with this particular band and so on. So it is important that we should actually destroy the secondary structure into the RNA species that kind of secondary structure is not present in the DNA because DNA is double standard. So we have to run the destructive or formamide gels actually.

So RNA sample and agarose gel contains the formaldehyde to denature the secondary structure present in the RNA and prevent the reformation of double standard region into the RNA structures. Now if I want to perform the formamide gels I require the specific materials and equipments and all the kind of things. So what are the materials you require? You require the agarose, you require the 10x MOPS buffer. So this is the composition of the 10x MOPS buffer and you also require the 37% formaldehyde and you require the RNA molecular weight ladders, you require the 0.5 molar sodium acetate, you require the 0.

5 microgram per ml ethidium bromide so that you can actually be able to use that for staining and you require the RNA free water. If you do not have the RNA free water then you can be able to utilize the standard procedure to prepare the RNA free water although RNA free water is readily available from the different types of vendors but if you do not have then you can actually be able to prepare the RNA free water in your lab. Then you also require the formamide and you also require the formaldehyde loading buffer with the composition of the formaldehyde loading buffer is that 1 millimolar EDTA pH 8, 0.5% bromophenol blue, 0.

25% xylene cyanol, 3% glycerol. Then you also require the loading buffer can be filter sterile by passing through a 0.22 micron filter and it can be allocated to the small volume and stored at minus degree. Then you also require the equipments such as you require

the autoclave gloves, you require the 50% water bath, you require the horizontal gel electrophoresis system, power supply, RNA-CD container for staining and destaining, you require a shaker, UV chambers, gel dock and flask for preparing the agarose gel. Now let us see what are the procedure. So in the procedure in the step 1 you are going to do the isolation of messenger RNA that we have already done with the help of the affinity chromatography.

In the step 2 you are going to prepare and cast the denaturating agarose gel. So the preparation of the RNA free water. So this is what I was trying to say that if you do not have the RNA free water then you can be able to prepare the RNA free water in your lab. What you are supposed to do is you dissolve the DEPC or diethyl pyrocarbonate in the deionized distilled water to a final concentration of 0.

1%. Remember that DEPC is a very strong inhibitor of the RNA. So if you add the DEPC it is actually going to kill the RNAs whatever RNAs is present into the water. And what is the source of RNAs in the water? The source of RNAs into the water is the different types of bacteria, small organisms and all that. So that has to be removed otherwise you will isolate the RNA and it is going to be degraded by these RNA species. Then you stir this solution in 12 hours at room temperature and then autoclaved to remove the DEPC.

So if you stir this solution for 12 hours in room temperature the DEPC what is present in this solution is going to kill the RNAs what is what is been present and then you can actually autoclave and autoclaving is actually going to deactivate the DEPC and then you can store this at room temperature. Now the third step is that step is the casting of the agarose gel. So in a flask add 1 gram of agarose to 75% RNAs free water. Heat the solution to melt the agarose and observe the disappearance of the agarose flecks. Then you allow the solution to cool down up to the 55 degree Celsius inside of fuming hood add 10 ml of 10X MOF buffer and 18 ml of 37 formaldehyde.

Remember that formaldehyde is a very corrosive liquid. So you should be very careful because it may actually cause the burning and other kind of accidents. Then we set up the casting tray with the comb and pour the gel into the fuming hood. That is why we remember that we are using the fuming hood because it is actually going to form the different very corrosive fumes and they may actually cause the you know irritant to the eyes and other kinds of face actually. So you once you add the 37% formaldehyde and then you are going to pour this solution into the casting tray.

I already written that formaldehyde is very toxic and can be easily absorbed through skin. So you wear the gloves and mask. Then you prepare the sample RNA sample for

running. So take the RNA sample and make up the volume to 6.

5 microlitre with the appropriate quantity. To each sample add 2.5 microlitre of 10X MOF buffer, 4.5 microlitre of 37 formaldehyde and 11.5 microlitre of formamide. Then you mix by vortexing and briefly spin to collect the sample at the bottom.

Then inside the hood add 5 microlitre of RNA loading buffer, mix by vortexing and briefly spin to collect the sample at the bottom. This is required so that you should not waste the sample actually. Then in the step 4 you are going to load the RNA. So fill the agarose denaturing gel prepared with the 10X MOF running buffers, load the RNA sample onto the lid.

So this is what it is showing right. You are going to load it with the help of the RNase free, DNase-RNase free tips right. And then you are going to run it for agarose gel. So in the step 5 you are going to run the denaturing agarose gel. So place the lid onto the buffer chamber and perform the electrophoresis at 5 volt per centimeter until the dye front reaches the two-third length of the gel. And then in the step 6 you are going to stain the agarose gel.

So in a RNA free container agarose gel is dipped into the 0.5 molar sodium acetate for 40 minutes at room temperature. You drain off the solution and dip the solution in 0.5 molar ammonium acetate containing 0.

5 microgram per ml ethidium bromide. Then incubate on room temperature for 30 to 40 minutes if required and the stain is too intense it can be destained by the 0.5 molar ammonium acetate for another 60 minutes. So initially you are going to dip the agarose gel into the 0.5 molar sodium ammonium acetate for 40 minutes at room temperature. Then you are actually going to stain it with the help of the 0.

5 molar ammonium acetate containing 0.5 microgram per ml ethidium bromide. So this is actually going to stain the RNA and then if the stain is too high then you can actually be able to use the destaining solution which is 0.5 molar ammonium acetate and if required the stain is too intense it can be destained by the 0.5 molar ammonium acetate for the 1 hour. Then transfer the gel to the UV chamber and capture the image with the gel documentation system and what you are going to see here is that you are going to if you run it along with the ladder that what you are going to see is the individual bands of the RNA species and you are also going to see the RNA what is being separated from your sample.

So you are going to have the two different types of samples and you can be able to get

separated. So this could be control sample, this could be treated sample. So that is how you are going to do. Now once you see that there is a very nice separation of the RNA then you can actually go move on to the next step of the northern blotting. So in the next step you are going to transfer the gel to the NC membrane and this step is exactly the same as what we have discussed in the southern blotting.

So place the gel in a RNA free petry dish rinse 4 times with the sufficient DNA's water soak the gel for 20 minutes in 0.05 normal NAOH and keep it in a enough 20X SSC for 30 minutes then you place the two pieces of what meant 1 mm paper of the size of gel on to the glass plate and wet it with the 20X SSC. Place the gel on the filter paper and remove any air bubble trapped between the gel and the filter paper cover slides with the plastic wrap. Once the membrane on the gel avoiding the air bubble flood the surface of the membrane with the 20X SSC and you place 4 sheet of what meant 3 mm filter paper of the same size on top of the membrane then you place a heap of water paper towel on top of the filter and add on the approximately 500 grams weight and leave for overnight. So what we are describing here is actually the capillary action transfer site that we have already discussed in the DNA in the southern blotting.

And then this is what it is going to be the procedure right. Now once the step 3 is done your RNA is going to be transferred on to the nitrocellulose membrane then you are going to have the next step. The next step is preparation of the radiolabelled probe or radiolabelled DNA right radiolabelled cDNA probe right. So preparation of probe DNA so preparation of the complementary DNA or cDNA multiple approaches are been developed to prepare the complementary DNA from the isolated messenger RNA in all approaches the 3 steps are performed. You are going to have the first strand synthesis with the help of reverse transcriptase then you are going to have the removal of RNA template and then you are going to have the synthesis of the second strand synthesis. So there are multiple method what people have used in previous when we were discussing about southern blotting we discussed about the random primer method and now we are going to discuss the another method how you can be able to prepare the cDNA probe corresponding to a messenger RNA and then you how you can be able to use that.

So this method what we are using is Gubber Hoffman method in which you are going to use the standard procedure first you are going to do a first strand synthesis with the help of reverse transcriptase followed by the removal of the RNA template and followed by the synthesis of the second strand. So in the Gubber Hoffman method in this approach after the first strand synthesis using the oligo dt primer in the presence of the reverse transcriptase and dntp's DNA RNA hybrid is treated with the RNA edge to produce the nick at multiple side then the RNA DNA polymerase is used to perform the DNA synthesis using the multiple fragment of the RNA as a primer to synthesize the new

DNA strand and this method produces the blunt and duplex DNA. So what will happen is that suppose this is the messenger RNA which you want to use for preparing the cDNA probe. So what you are going to do is you are going to add the poly dt primers and using these primers the reverse transcriptase is going to make the RNA.

So this is going to be RNA and this is going to be DNA. This means at the end of the reverse transcriptase reactions you are going to have the RNA and DNA as the hybrid form. So you are going to have the hybrid species and these hybrid species are actually been targeted by the one RNA which is called as RNA's edge. So RNA's edge is very specific for those RNA's which are been part of this hybrid species. So when you treat this with the RNA's edge and it is actually going to make the cleavage. So it is going to cleave this fragment into the multiple places and because of that these small fragments are going to be attached but they are actually going to have the nicks.

So these are the nicks what you are going to be present and now what you are going to do is you are going to add the DNA polymerase. So what DNA polymerase is going to do remember that we have already discussed about the replication. So what DNA polymerase is going to do is it is going to utilize this as a primer and then it is actually going to start synthesizing the whole DNA and as a result you are going to get the cDNA of this corresponding messenger RNA. So it is going to have the double standard DNA. And then the second step is the radio labeling of the cDNA probe what you have prepared.

For that you can actually use the terminal transferase method. So in this method a terminal transferase enzyme will be labeled the probe at the end to the last nucleotide of the probe and the probe is incubated with the labeled nucleotide and terminal transferase enzyme will add the labeled nucleotide at the end. A partial purification with gel filtration column will give you the labeled primers. So then this what we have is you have this is the double standard cDNA what you have prepared and then what you are going to do is you are going to treat this with the alpha exonuclease and that is actually going to create the nicks and then what you are going to use you are going to use the terminal transferase with the help of the radioactive nucleotides and then what will happen is that you are actually going to add the nucleotide on both the sides and that is how you are going to have the radio labeled probe and now you can actually be able to use these double standard DNA. So you are going to actually this you know you are going to denature this and as a result you are going to have the two strands you are going to have two strands from this and that is how you when we are able to use this as a probe for subsequent steps of hybridizations.

And in the step 5 you are going to have the hybridization. So rinse the membrane in 5x

SSC then place it on the sheet of the Whatman filter paper heat it for 2 minutes in full power in a microwave oven and further cross linked with the UV rays for an appropriate time. So you can use the UV rays of 250 nanometer wavelength. So why it is important that you are going to do the cross linking because you do not want the disturbance of that appearance. Then for pre hybridization of the membrane wet it into the 5x SSC next place it in a pre hybridization solution kept in a tube and incubate in an incubator with the rotation at 60 degree without probe. Remember that all most of these steps we already discussed that you are actually requiring a hybridization chamber and hybridization chamber has a tube like this and in this tube this tube will go into a hybridization chamber.

And then in this you can actually be able to place your membrane and then you can just fill this solution and put it into the hybridization chamber and that will maintain a temperature of 60 degree Celsius and then it is actually going to keep rotating. And because of that the hybridization buffers and everything will keep interacting with the membrane. So for hybridization denature a double standard probe by heating it into a water bath for 10 minutes at 100 degree Celsius and then by quick chilling. Then pipetting the desired volume of the probe into the hybridization tube and incubate overnight with rotation at 42 degree Celsius for DNA probe.

The membrane was washed 5 minutes in 1x SSC 0.1% SDS at room temperature and washed for 5 minutes at 68 degree Celsius in 0.5% SSC and 0.1% SDS. Air dry the membrane exposed to the sensitive x-ray film at minus 80 degree Celsius with intensify the screen prepare an autoradiogram after developing a film.

Now the next step is the autoradiogram. So you are going to do the air dry the membrane exposed to the sensitive x-ray film at minus 80 degree Celsius with intensify the screen prepare an autoradiogram after developing a film. So why we do at low temperature because we just do not want the spreading of the radio activity. You just want to keep the signal very intensified so that when you are doing this actually it is going to give you a band instead of giving you a blob. If you do it on room temperature or any other temperature then the radioactive will scatter actually and it is going to give you the blob rather than a band actually. So just to reduce the activity and that is why we are doing this like 48 hours or 72 hours exposure because you are increasing the exposure time but you are reducing the activity of those beta particle meeting atoms so that the band what you are going to see will be very sharp.

Now the intensity of the band on the developing pale which is a measure of the specific messenger RNA in the sample can be compared in the different samples. So what it is actually going to tell you it is actually going to tell you the expression of a gene or I will

say transcriptional activity. Transcriptional activity of the gene or the target gene whatever you are probing. If you compare this if you have two samples suppose you have a control sample you have a treated sample and suppose this is the band what you get from the control sample and this is the band what you get in the treated sample or vice versa then it will going to say you that there is a enhanced production of the messenger RNA or of that particular gene under this particular treatment. That indirectly mean that it is actually enhancing the expression of the gene although you have to verify that with the help of the other techniques such as western blotting and other kind of techniques to verify that whether the transcriptional activity what you observe for that particular gene is also been reflected in terms of the translation or not.

Because sometime it happens that you have the RNA species but it may not be get completely been translated. So sometime you may see a increase in transcriptional activity but that may not be into the translation. So that kind of variation could be possible and that is why you are supposed to do the further verification with the help of a translation into the translation with the help of another technique which is called as the western blotting. So this is what we have discussed so far we have discussed about southern blotting we have discussed about the northern blotting and as you recall while we were discussing about the northern blotting it is actually going to tell you about the transcriptional activity of a gene. But if you want to confirm whether that transcriptional activity is also been resulted into the translation of that particular gene fragment and giving the protein or not then you are supposed to do the next blotting technique and that first blotting technique is called as the western blotting and that we are going to discuss in our subsequent lecture.

So in our today's lecture what we have discussed we have discussed about the northern blotting we have discussed about the different steps of the northern blotting and how you can be able to separate the messenger RNA how you can be able to isolate the messenger RNA with the help of the affinity chromatography how you can be able to separate the messenger RNA onto the denaturating formamide gels.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 08
Molecular Techniques (Part 1)
Lecture-35 Western Blotting (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering, IIT Guwahati. So, in that we have discussed about the apoptosis cell divisions and the autophagy. So, subsequent to that we have also discussed about the central dogma of molecular biology and within the central dogma of molecular biology we discussed about the replication, transcription and translations and we have discussed about all of these events in the prokaryotic as well as the eukaryotic system. We have discussed about how these events are being regulated within the cell with the help of different types of modifications and other kinds of regulation at the cellular signaling levels and all that. Subsequent to that we have also discussed about the different types of techniques what is being evolved or what is being developed using the molecular biology principles and in this series we are in the current module we are discussing about the different types of blotting techniques. So, far what we have discussed we have discussed about the southern blotting where and the purpose of the southern blotting is to detect the DNA fragment into genome and then following that we have also discussed about the northern blotting.

So, if you recall in the previous lecture we have discussed about the applications of the northern blotting and how the northern blotting is being performed. So, when you want to perform the northern blotting you are supposed to first isolate the total messenger RNA from the cell and then it has to be separated on to a denaturing agarose gel and following that we have also discussed about the transferring of these RNA molecules on to the nitrocellulose membrane following that you have to hybridize that with the suitable radioactive probe and then you have to develop that with the help of the autoradiogram and then at the end you are going to get the pattern and this pattern is actually going to tell you about the transcriptional activity of that particular gene fragment and it is going to tell you how it is modulating in the different types of response. Now in today's lecture we are going to discuss about the another technique which is called the western blotting and the purpose of the western blotting is to detect the proteins on to the cell lysate or into the cellular system and protein detection of protein is very important because it is going to give you the idea about the translational activity of that particular gene or gene product and the detection of western blotting or detection of protein with the help of western blotting is a multi step process where initially you are going to run the you are going to separate the proteins on to a

SDS page following that you are supposed to transfer that on to a nitrocellulose membrane and then you are supposed to treat that with the primary antibody followed by the secondary antibody and then followed by the development of the depending upon the secondary antibody are supposed to use the suitable substrate to give you the bat. So what you are going to see here is that you are going to first run the so this is the step number 1 that you are going to run the proteins on to a SDS page then you are going to transfer that on to a nitrocellulose membrane and then you are going to do so this is the step number 2 and then you are going to do a blocking step followed by the primary and after the primary you are going to have a blocking you are going to have a washing step so that the excess antibody can be removed and then so this is going to be the step number 3 this is going to be step number 4 and then you are going to have the secondary antibody followed by the washing and then you are going to add the substrate or the different types of reagents what is required for developing and then it is actually going to give you a band corresponding to a particular protein. So what you see here is that it is mainly having the 5 different steps step number 1 you are going to resolve the protein mixture on to on SDS page.

In step 2 you are going to transfer the protein band on nitrocellulose membrane just like as we have discussed about the southern as well as northern blotting step 3 you are going to treat this with primary antibody step 4 you are going to do a washing and you are going to treat that with the secondary antibody and in the step 5 you are going to do a washing and you are going to develop that with the help of develop by the substrate or you if some cases you are using the radioactivity then you are going to use the autoradogram. So if you see the technique you can be able to very clearly see that it can be divided into 2 parts one part where you are going to run the protein on to SDS page and then you are actually going to do the subsequent steps by after transferring the protein band on to the nitrocellulose membrane. So just for the convenience of understanding this we are first going to discuss about the protein resolution resolving of the protein mixture on to the SDS page following that we are also going to discuss about the all these steps in a subsequent lecture. So the western blot is western blotting is a popular technique to detect the specific protein present into the crude lysate or the homogenate it uses the separation of the protein different proteins into the gel electrophoresis which is called as the SDS page and then the transfer of the protein on to a solid support such as nitrocellulose membrane a primary antibody direct against the protein of interest and the secondary antibody is used to detect the primary antibody and give either colored or the chemiluminescent substrate. So what we are going to do is we are going to first understand so it is actually having the 2 part one you are going to understand about the electrophoresis so that you can be able to understand how the you can be able to separate the so protein runs on the SDS page so how you can be able to separate the protein on to a SDS page and the second part is the transfer of protein band

and following by the subsequent treatments so that anyway we have discussed.

So let us first discuss about the electrophoresis and we are not going to discuss in detail about the electrophoresis we are only going to tell you about how to perform the electrophoresis and then we can use that technique to tell you about how to perform the electrophoresis and how to perform the SDS page so that it will become a complete learning experience. So what is electrophoresis? Electrophoresis is that you are going to run a charged particle into a electric field so you can imagine that if you have a charged particle with Q then it is actually going to have this process is called as the electrophoresis. So electrophoresis is an electrokinetic process which separates the charged particle in a fluid using a field of electric charge it is most often used in the life science to separate the protein molecules or the DNA molecule and can be achieved through the several different procedures depending upon the size and the type and size of the molecule. So you can imagine that if you have a charge Q which is been resolved and it is so it is actually going to experience electric field or electric charge which is electric force which is called as F is equal to Qe and it is also going to experience a retardation forces and these retardation forces are going to be dependent on to the radius of this particular molecule and as well as the viscosity of this particular media and that is the friction forces is going to be called as $6\pi\eta RV$ and where this molecule is going to stop the place where the F is always equal to the friction forces and at that point the electrophoretic mobility is going to be proportional to the Ze by the $6\pi\eta R$. So Ze is actually and so if you if you subsequent if you further simplify this what you will see is that the electrophoretic mobility is directly proportional to the E by R actually and that means it is actually going to be proportional to the charge by mass because R is directly proportional to the M because for most of the spherical molecules R is proportional to the M and E is proportional to the charge.

So it is going to be Ze is the charge actually. So hence the electrophoretic mobility V is directly proportional to the charge and the inversely proportional to the viscosity of the media size and the shape of the molecule in case of relative mobility it is directly proportional to the charge by radius of the molecule for a globular protein the radius of the molecule is related to the molecular mass of the molecule and that is why the relative electrophoretic mobility is charged by mass. So you can actually be able to run the electrophoresis in two different mode the mode in which the charge and mass both are actually going to be intact and that mode is called as the native page. Other mode is that you if you neutralize the charge what is present onto the protein and then it is going to be called as the SDS page and the difference between the SDS page and the native page is that in the native page the electrophoretic mobility is directly proportional to the charge by mass whereas the electrophoretic mobility in the case of SDS page is going to be inversely proportional to the mass because you have already you have

already nullified the charge. So this is very very important because it is actually going to separate out the molecule based on the mass and that is how you can be able to very precisely predict the mass of that particular protein if you look at the pattern. So what we are discussing in this is about the SDS page we are not discussing about the native page because the most of the western blotting procedures does not involve the running of the native page.

Now what are the things you required for running the electrophoresis? So you require the following components you require the gel cassette you require the electrophoretic chamber you require the cords you require the power pack or the instrument which actually can supply the requisite power supply then you also require the electrophoretic tank and then you also require the comb so that you can be able to prepare the wells. What are the chemicals you require? So you require the different types of reagents in preparing the gel electrophoresis you require the buffers and reagent for the electrophoresis you require the NN, NN tetramethyl ethylene diamine which is also in short form it is called as TEMID so it catalyses the acrylamide polymerization then you require the ammonium bar sulfate or APS it is an initiator for the acrylamide polymerizations then you also require the TrisSCL so it is the component of the running and the gel casting buffers then you require the glycine and it is the component of the running buffer then you require the bromophenol blue and it is the tracking dye to monitor the progression of the gel electrophoresis then you require the comasibilian blue R250 it is used to stain the acrylamide gel you require the SDS or sodium dodecyl sulfate it is used to denature and provide negative charge to the protein and it also require the acrylamide and acrylamide is a monomeric unit to prepare the gel so you always have the combination of the acrylamide and the bisacrylamide and you also require the bisacrylamide and that is the cross linker for the polymerization of acrylamide monomer to the form gel how the bisacrylamide and acrylamide come together and form the network or the mesh so that you can be able to have the retard you can you are going to have the frictions and that actually is been responsible for the separation of the different molecules so you are going to have the acrylamide polymerization there is a complete mechanism in which the acrylamide is and bisacrylamide is actually going to be mixed up with the tamid and APS and that is how it is actually going to make the cross link so the bisacrylamide monomers are actually going to make the cross link these fibers and that is how it is actually going to make a network. So ammonium bar sulfate in the presence of tamid form the oxygen free radicals and induce the polymerization of acrylamide monomer to form a linear polymer these linear polymers are interconnected or connected by the cross linking with the bisacrylamide monomer to form the 3d mesh with the pores the size of the pore is controlled by the concentration of the acrylamide and amount of bisacrylamide in the gel in a vertical gel electrophoresis system we cast two different types of gel stacking gel

and the resolving gel for the resolving gel solution is prepared and poured into the gel cassette for polymerization a thin layer of organic solvent is attached or layered to the top the entry of oxygen and this is been done so that you can actually be about avoid the entry of oxygen and as well as to make the top surface smooth oxygen is actually inhibitor of the free radicals and if you allow the oxygen entry then it is actually going to destroy or it is actually going to inhibit the acrylamide polymerization. So it is actually going to inhibit the acrylamide cross linking or polymerization because acrylamide polymerization is associated with the free radical formation and that is can be stopped after the polymerization of the resolving gel a stacking gel is poured and the comb is fitted into the gel for construction of the different lanes for the samples. How you are going to run the gel electrophoresis? So gel electrophoresis you are going to have the two different types of gels you are going to have the resolving gel and you are going to have a stacking gel the purpose of both the gels are very different the stacking gel is required for the stacking of the sample whereas, the resolving gel is required for the separation of the molecules.

So first the resolving gel is prepared and poured into the gel cassette for polymerization and thin layer of organic solvent is layered and all that and then you are actually going to have the separating the pouring of the stacking gel. So there is a recipe available to which you can be able to prepare the solution for the resolving and as well as the stacking gels and these are the different components what is required for running the electrophoresis system. Now the question comes how the stacking gel and what is the relevance of the stacking gel into the gel electrophoresis. So the sample is prepared in the loading dye containing the SDS, beta mercaptoethanol, the glycerol to denature the sample and preference of the itself loading. So you can imagine that you have a well and in which the samples are filled like this.

So if you fill the sample like this some samples are going to be present here and some samples are going to be present here just like in a race some runners are present here and another runner is present here. So there is a difference between these two runners and that is why you see very clearly that this runner is actually running in a circle which is of a larger diameter whereas this guy is running in a smaller diameter. This means the distance what they are going to run is actually be the same. So same is the concept when you are want to do the stacking because what you have is you have a well and this well is standing like this. So you are actually going to have a sample here and you are going to have a sample here.

This means these two guys are even if they are of similar molecular weight they are separated from each other to this distance and that is why if you do not do the stacking this guy which is like the number 1 and this guy is number p. Number 1 guy is always

going to be remain on the behind side. So number 1 will come here and the p will come here because even if they are of the same molecular weight this means you are supposed to device a mechanism so that the 1 is also going to come here and then the 1 and p also should come together and they will run together so that they are actually going to show you the real separation and real separation would depend on to the molecular weight. So as the samples are filled vertically there is a distance drift between the molecule at the top versus the bottom lane and the problem is taken care once the problem runs to the stacking gel. So stacking is and that is why you are actually going to have a specific composition of the stacking gel.

So what is the composition of a stacking gel? In the stacking gel what you are having is they are actually going to have the this SCL of pH 6.8 and that is a very very important to take care of this problem. So what you are going to do is the problem is been taken care once the sample runs to the stacking gel the pH of the stacking gel is 6 point at and this pH the glycine is moving slowly in the front whereas the tris glycine is moving fast as a result the sample get sandwiched between the glycine tris and get stacked in the form of a thin band. As the sample enters the resolving gel with the pH 8.8 the glycine is now charged it moves fast and now sample run as well as their molecular weight.

After tracking dye reaches to the bottom of the gel is taken out from the glass plate with the help of the spatula and it is stained with the chromatic blue and blue R250 the dals type present on to the gel. So this is what is actually going to happen. So this is actually going to be the well and you are going to have the sample which is on the top and you are also going to have a sample at the bottom.

So this is and you are actually going to have a pH of 6.8 in the stacking gel. So at pH 6.8 the glycine which is also been present in the running buffer is actually going to work as a block. So it is not going to allow these molecules to run beyond this. And on the top you are actually going to have tris ions and these tris ions are actually going to run from the top. So this is actually going to put a pressure on to these molecules and that is how it they will come they will they are free to move but these guys are not free to move because there is a block there is a glycine block at the front.

And as a result what will happen is that when they are running they are actually going to come closer closer closer like that and at the end what will happen is that they all will be at one place they all will be at one place and this is going to be called as the stacking. And as soon as this is been done you are actually going to have the resolving gel. So this is going to be a resolving gel and they will enter into the resolving gel and then the resolving gel has a pH 8.

8. So once the pH 8.8 is there this glycine block is actually going to be removed and then the molecule will run as per their molecular weight. So they are actually going to be and then the V is inversely proportional to the 1 by m. And that is how the larger molecule will run slower and the smaller molecule will run faster. So just to and then once you are done with the once and how long the gel is running for that purpose you are actually adding a tracking dye which is and then the when the tracking dye is going to reach at the bottom of the gel then you can actually be able to remove this particular thing and then you can if you are doing the western blotting then you can just keep it as such. But if you are doing it for the just for looking at the pattern and then you can actually be able to stain it with the comasibilian blue and that is how you are going to get the pattern.

So you are going to see a pattern of the proteins after the staining. So just to explain you all these steps and how to make you familiarize with the whole process we have prepared a demo video where we have taken we have shown you how to assemble the acids, how to cast the resolving gel, how to cast the stacking gel and so on and it will help you to understand the whole process. Hi everyone myself Suram Banish research scholar at department of biosciences and bioengineering at IIT, Guwahati. In this video we will demonstrate you how to run a SDSS gel and how to prepare various reagents required for the running of SDSS gel and what are the different instruments we can use. So here this is the gel casting stand so where we can use this glassware to prepare the gel. In between there is a space where we can pour our gel solution then we will keep for some time at least 20 to 30 minutes let it solidify then we will prepare stacking gel then we will load the protein solution.

So here before doing that we need some reagents. So what are those reagents? The first reagent we need for this experiment is acrylamide. So generally we will prepare acrylamide 30 percentage. 30 percentage means 29 grams of acrylamide and 1 gram of bisacrylamide. This both we can use 29 is to 1 ratio in 100 ml of water to get 30 percentage of acrylamide.

So both these are neurotoxic so we have to wear gloves always. After this we have to prepare resolving gel. For resolving gel we need 1.5 molar Tris HCL pH 8.8. In addition to that we also need 10 percentage SDSS prepared in double distilled water and also 10 percentage ammonium per sulphate and also TEMED.

The role of ammonium per sulphate and TEMED we can see during preparation of gel. They acts as a catalyst. After solidifying we have to use we have to prepare stacking gel. So stacking gel is nothing but composition is same but we can say it is diluted.

It contains pH 6.8 Tris HCL and remaining components same but in less quantities. So after preparing the gel we load the marker and the protein which is denatured at 100 degree Celsius for 3 minutes. After that we will fix this gel into this one. We will keep in this reservoir then we will connect to the power pack and run the gel. So this is the overall introduction of how to prepare a SDSS prepared gel.

So let us start with preparing gel. We will learn more things while preparing the gel. Before preparing the resolving gel we have to prepare the casting set up the casting gel.

So this is the glass plate. This is very thin one. So this is the main glass plate. This is 1.5 mm glass plate. It is available in 1 mm glass plate also. If your loading solution is less like you want to load only 20 micro litre, 30 micro litre then 1 mm gel is good enough.

But if you have extended volumes like 70 micro litre you can use 1.5 mm. You have to arrange like this short plates on this plate and the bottoms should be equal. Then we have to put in this one, this tray. Then we are going to keep like this. So we have to check if we perfectly set up this one then there should not be any leakage.

But if there is any leakage your resolving gel may leak out and you will get nothing. So in that case we have to check it prior to pouring the gel. So whether it is okay or not. So I am going to use milk water. After checking the gel if there are any leakages or not.

So we move forward for preparing the resolving gel. So the composition is given in this slide.
Please go through that slide.

This is just water. First I used water. I am going to add sequentially 4 ml of water. Now I have to add 3.3 ml of already prepared 30 percent acrylamide. Already in introduction I explained how much percentage we have to prepare and how much quantities of acrylamide and bisacrylamide need to take.

So here we have to add 3.3 ml of acrylamide solution 30 percentage.

So I have to adjust 300 micro litre. The next component is 1.5 molar Tris PH 8.8. We have to add 2.5 ml. Next component is SDS. Here SDS functions as place as dual role. Like one thing is that it gives negative charge gross negative charge on the polypeptide chain. The next component we have to add is SDS 10 percentage SDS.

We have to add 100 micro litre of SDS to resolving gel. It plays very crucial role in polyacrylamide gel electrophoresis. Like it imparts negative charge on the polypeptide

chain so that despite of their charge they will move based on the molecular weight. So I am going to add SDS. The other important thing is that 10 percentage ammonium persulfate. Ammonium persulfate which is catalyzed by the timid provides free radical species which accelerate the forming mesh like say in acrylamide gel.

Like it will catalyze forming the mesh. So this is the 10 percentage APS. I just add 100 micro litre of 10 percentage APS to resolving gel. In final step we have to add timid. Timid after finishing after adding all the components at the end of the gel we have to add timid.

Because if you add earlier it will quickly facilitate the polymerization. So you cannot take out with the pipette. So it completely solidifies. So that is why you have to add at the end of the cell. So I am going to add 5 micro litre of this timid which catalyzes the ammonium persulfate. Ammonium persulfate in turn provides free radical species and free radical species accelerate the polymerization.

This is the overall principle of this resolving gel. So I will add timid. We have to mix properly. Then add slowly at one corner. So after this we have to overlay with on the top layer we have to overlay with some solvent like 2 butanol or isopropanol or with water. So why we are doing this? Because if the gel is exposed to air then the oxygen from the air will interfere in the polymerization of the gel.

So we have to add either water or 2 butanol for this purpose. Now we have to check whether it is solidified or not. So it is solidified. Now we have to remove the overlay layer like we have used water. So no need to remove. If you are using isopropanol or butanol you have to remove that and wash with the milky water.

So now we will start preparing the stacking gel. The compositions are given in the video. We have to add 3.4 ml of water first. Next 830 microlitre of acrylamide.

630 microlitre of Tris-HCL pH 6.8. 50 microlitre of imide and 50 microlitre of SDS we have to add. At the end we have to add 5 microlitre of Tb. We have to mix properly after adding the Tb. Now we will wait until the gel gets solidified. Then we shift to the buffer tank and then we will run the gel.

While the stacking gel is solidifying we have to prepare sample for loading the SDS-plane gel. So for that we have to prepare loading gel 10x or 6x loading gel. It mainly contains 250 ml of millimolar Tris-HCL pH 6.8, 30 percentage glycerol, 10 percentage SDS and 0.

0.05 percentage of bromophenol. So here we can add 10 millimolar of DTT as a reducing agent. SDS mainly works as imparting negative charge on the polypeptide chain and DTT reacts on the disulfide parts. If you have dimer which you can see as a monomer in SDS-plane gel. Suppose you have 20kDa, 20kDa that means 40kDa protein which is a dimer actually. You can see only 20kDa band corresponding to that protein because DTT reacts on the disulfide part and you can see only single band.

If you want to see actual molecular weight you have to run it on native phase where there is no reducing agent or no detergent. The other thing is glycerol. While loading the gel since the protein solution is not that much dense it may come out from well. So in order to prevent this thing we have to load with the denser solution like glycerol. So 30 to 50 percentage glycerol is sufficient for keeping the protein solution intact in the bottom of the well.

So other thing bromophenol blue. Bromophenol blue we use for just tracking the how much gel completed. So this is the loading time. So we have to take the protein solution. Here we already prepared 10 percentage of loading time. So that means this is 10x loading time.

So we have to prepare 1x to mix with the protein solution. So this is 100 ml of solution, loading solution. We mix 10 microlitre of loading time to this protein solution. You can tap down or pipette this protein solution. Then we have to heat it for 3 minutes at 100 degree Celsius so that all the polypeptide chains, I mean dimers are present. If any multimers are present they will break down and you can see nice band.

So I am going to heat this at 100 degree Celsius for 3 minutes. This is the remaining of stacking gel solution. So we can see it is solidified. So that means the stacking gel also got solidified. We have to remove that gel and fix it into this one and we have to keep inside the tank.

So just take out the gel. So inside this tank we only have this side one. You have to cover other side also. So for that we use dummy plate. Just hold it tight and close this thing.

After that gradually adjust the gel length. So just we have to fix like this. Once fixing here we have to add this running buffer. The running buffer contains 15 grams of tris, 72 grams of glycine and 5 grams of SDS for 2 liters of solution, 1X solution. So this is 1X I have already prepared. I am going to add. We added in this tank but the main tank surrounding to this one we have to add up to the mark.

So for reference you can see here for 4 gels we have to add till here the buffer we have to load outside this gel. So for 2 gels here for 1 gel we can add like this. This is the power pack where we can adjust the how many volts we want to run.

The protein samples are ready. We heated sufficient time. Now we have to load this. So we have to remove the column carefully. First I am going to load marker or protein ladle.

Next I will load sand.

Once the loading goes over we have to fix this gas tank. I am going to set it at 170 volts. This thing has Time supply point 0, lightAST. So, we will try to heat it with the staining solution which contains pumasi brilliant blue and along with methanol and water. So, then we will try to de-stain with the water by heating. But in another way, the simplest way is we will just stain the engine for 2 hours, then we will de-stain overnight. So, I am going to show the simplest way.

First we will stain in pumasi brilliant blue staining solution, then we will de-stain in methanol water containing salt. So, I am going to start the children, then I will remove it and show you how to remove the gel. Take out the glass piece. And here we have to be very careful while taking out gel, otherwise the short plates may grow. On a corner we have to take and lift the gel like this.

So, keep the gel in a staining box which is more or less a plastic one but it can sustain the gel. So, then I am going to add staining solution. I will keep it for a rotation on a shaker for at least 2 hours, then we will de-stain the solution. So, once the time is over, after 2 hours we will de-stain the solution. Now, we kept 2 hours in staining solution.

As we can see the staining is over, like we can see the gel completely turned into blue. So, we remove the solution. Then I am going to add de-staining solution. And I will keep this on a rocker for 2 hours for de-staining. So, the composition contains for 100 ml of de-staining solution, 40 ml of water, double distilled water and 40 ml of methanol and 10 ml of glacial acid.

So, I am going to keep this on a rocker. We have run the gel and stained and de-stained. Now, we will capture the gel image. So, you can see manually also, but for record purpose, we have to capture it through gel dark. So, this is the gel dark imaging system from Bio-RAD.

So, I will show you how to take the capture the images. So, here we will use white tray. There is another one, grey or UV tray is also there. So, there you can see any fluorescent one or stained with the ethidium bromide or blots, chemiluminescent blots, you can use

that. But for normal protein imaging, we can use this white tray.

So, you have to open properly. This is very important step, you have to align the tray in a proper way. So, otherwise it will show error. So, once it is over, you just push it back. So, we have to log on to account. So, this is SDSS gel, you can select the application whatever you want. So, here nucleic acids, protein gels, blots, three different categories are there.

So, we are observing here protein gels, protein gels stained with the kumasi blue or white tray, we are using white tray. So, this is the right tray. You can use kumasi blue stained one grey tray also, but we are using as we are using white tray, so we will use kumasi blue. So, auto optimal, then I will ask for capture. So, it will take one to three minutes based on the signal intensity.

So as we can see, it is optimizing the signal intensity, you can minimize this one also. So that you can see the gel image. So, now it is over. If you want to do any modifications to images, for suppose you want to decrease or increase the signal intensity. So, this kind of changes you can do. So if you want to send this gel, you can have send itself. If you have any drive connected to this one, you can send directly to that one, that thing.

So for image analysis part, we will show in the upcoming video, how to analyze the what this band of interest correspond to which molecular weight. So we already loaded the molecular weight one, so we can easily find out using image lab software. In this video, we have learned that how to prepare a SDSS gel and how to run it, what are the precautions need to be taken while preparing the gel and how to record the gel using gel documentation system. So I hope this will give you a list of how to prepare and run a SDSS gel and analyze the protein sample. So this is the part one of the western blotting where we have discussed about the how to resolve the sample on to the SDSS page. And what we have discussed, we have discussed about the basic principle of the electrophoresis, we have discussed about the different component what is required for the running the electrophoresis and then we also taken the crucial mechanistic and as well as the technical steps what is required for running the SDSS page.

And at the end, we have also shown you a demo video how to prepare the how to prepare how to cast the gel, how to run the gel and how to stain the gels. So with this, I would like to conclude my lecture here in our subsequent lecture we are going to discuss some more aspects related to western blotting. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 08
Molecular Techniques (Part 1)
Lecture-36 Western Blotting (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from Department of biosciences and bioengineering IIT Guwahati and what we were discussing we were discussing about the different aspects of the western blotting in this particular lecture. And so far what we have discussed we have discussed about how you can be able to separate the molecule based on the molecular weight in the technique which is called as the SDS page. So, we have discussed about the how to prepare the gel, how to cast the gels, how to run the gel, how to perform the stacking and so on. And then at the end we have also discussed about how to stain the gels and how you can be able to visualize the pattern.

If you want to use these gels for western blotting then you are not supposed to stain these gels instead you can actually be put it into the transfer. So, now in the this particular lecture we are going to discuss about how to perform the western blotting. So, now we assume that you are stacking your gel is ready with the protein resolved and then you can be able to use these gels for performing the western blotting. So, in a western blotting what you have is so, is a step 1 you have are going to resolve the sample onto SDS page.

So, that you are going to see a pattern of the gel and now you are going what you are going to do is you are going to transfer these onto a nitrocellulose membrane. And then you are actually going to do a primary treatment, secondary treatment and ultimately you are going to use the suitable substrate to see the bands. Now let us talk about the step number 1. So, step number 1 you have already we have already discussed in our previous lecture that how you can be able to resolve the sample onto the SDS page. Now we will just start discussing about the step number 2.

So, for step number 2 onwards what are the materials required you are actually going to require the following materials you require E. coli cell expressing GFP. So, this is the sample actually this is the sample what you require for doing the western blotting. Then you require the protein standard markers, you require the transfer buffers, you require a transfer membrane such as 0.45 micro meter nitrocellulose membrane or the PVDF membrane, you require a plastic tray, you require spatula, you are require the blotting sheets.

So, 3 mm thick cellulose blotting sheets, then you require the semi dry electro blotting technique units, then you require reagents for forming the electrophoresis, you require the NTGFP antibody that is the primary antibodies and then you require the anti rabbit IgGHRP and then you also require the developing reagents such as the chemi reagents or the western blotting. So, as far as the procedure is concerned in the step 1 you are going to prepare the samples. So, preparation of the sample depends on the sample type for the tissue for solid tissue such as tumor or whole liver or brain it is first mechanically been broken down into the individual cell using a blender, homogenizer or by sonication once the individual cells are obtained it is been processed as described. For cells individual cells are incubated with the lysis buffer containing detergent along with the protease and phosphatase in a butter cocktail to protect the sample from the degradations. And then the step 2 you are going to do the electrophoresis of the sample.

So, this anyway we have discussed in the previous lecture. Then in the step 3 you are going to do the transfer of the protein gel on to the blotting membranes. So, this can be done 2 ways one is in 2 steps first is the preparation of the transfer membrane and the second is the assembly of transfer operators. So, cut the membrane of the same size as gel right. So, suppose you have a gel of this size then you can actually be able to cut the membrane of the same size or slightly bigger size actually because the nitrocellulose membranes are costlier.

So, you should be little careful with the size because you should not take very big so that you are going to waste. Then here you have the variations if you are using the nitrocellulose membrane the place the membrane the transfer buffer and observe that the liquid was wicking the wick the membrane areas appear as white spot need the special intentions right. So, when you are actually going to soak these membrane into the transfer buffer what you are going to see is the some portion is going to be remain dry. So, it has to be submerged completely so that it should not show you any white patches it should just show you the clean membrane. Then for the PVDF membrane it actually going to be require additional step of charging so that you should be able to convert the PVDF membrane because the PVDF membrane or polyvinyl dihydro fluoride membranes are hydrophobic in nature.

So, they will not going to bind the protein molecules instead you are suppose to convert or impart a charge on to the membrane so that they will be actually going to attract the protein molecules. So, we that you are going to do is by immersing the membrane into a 100 percent ethanol for 15 to 20 minutes. So, when you immerse the ethanol membrane into a 100 percent methanol it is actually going to give you the it is going to react with the membrane and it is going to give a polar environment. Then you decant the membrane methanol and submerge the membrane into a transfer buffer for additional 10

to 30 minutes and that is actually going to do the same is going to make the membrane suitable for transfer activities. Then the you are going to do the assembly of the transferic acids.

So, remove the stacking gel from the page and equilibrate the gel into a transfer buffer for 10 to 15 minutes. This means you are going to have a stacking gel. So, you are going to have a stacking gel. So, what you are suppose to do because the stacking gel is not required you can just cut this stacking gel and you can only use the resolving gel because the protein bands are only going to be present on to a stacking gel. So, unnecessarily why you should use this portion because otherwise you are suppose to waste the some amount of nitrocellulose membrane.

Then pair place a pair of blotting sheets which are already been soaked into the transfer buffer on to the anodic plate usually this plate is black in color. So, just keep the two one pair of the blotting sheets on to the black plate. Then you place the transfer membrane on top of the blotting sheets and remove the trapped air by rolling the test tube or the glass rod. So, you are going to have a glass rod and that glass rod you should roll it on to the on to this membrane. So, that you all the trapped air between the blotting sheets and the membrane can be removed.

Then you place the gel on top of the membrane and gently remove the trapped air bubble by rolling the test tube or the glass rod. Then you place the another blotting sheets already saturated with the transfer buffer on top of the on the top and remove the trapped air by rolling the test tube or the glass rod. And finally, keep the cathode rate usually the red color and tight the transfer buffer by the four screws. So, you are going to actually having a four screws on top of the four corners. Remember when you are doing the screwing it should be a longitudinally which means if you are doing it 1, 2, 3, 4 your 1 and 3 should be screwed together and 2 and 4 should be screwed together.

So, that you should not have the unbalanced you should not have a any kind of you know the any kind of pressure into this. Because if you do it 1 and 2 then your this portion is actually going to be having a some kind of pressure and as a result it may actually take up some air. Then you are going to transfer the protein from the gel to the membrane. So, place the transfer cassette into the tank filled with the transfer buffer connect the transfer operator to the power supply unit and apply constant voltage for 1 hour after the transfer disassemble the whole cassette and carefully remove the transfer membrane and check the protein transfer by the one should use a pencil and label the different lengths and then you are going to have the step number 4. The step number 4 you are going to do the blocking.

So, wash the membrane with distilled water to remove any remaining poncho stain put the membrane in blocking buffer containing the 5 percent skim milk or 5 percent BSA for the detection of the phospho proteins. So, if you are doing it for the normal protein then you can actually be able to use the 5 percent skim milk as a blocking agent, but if you are doing it for the specific proteins such as the phospho proteins or the glycoproteins and other kinds of proteins then you should do use the BSA because then because the skim milk has lot of alkaline phosphatase. So, that may actually you know destroy the or damage the samples. Then you have the probing. So, in western blotting probing can be done in 2 ways a 2 step probing and the 1 step probing.

So, in a 2 step probing in 2 step probing scheme the membrane is first probed with the primary antibody to recognize the protein of interest. So, and then membrane is probed with the primary antibody with an appropriate dilution for 1 hour at room temperature membrane is washed with buffer containing non-ionic detergent titronics 100 and reprobed with another antibody directed against the primary antibody. The secondary antibody is coupled with an enzyme either HRP or the alkaline phosphatase or a fluorescent dye. Washed membrane is incubated with another secondary antibody with an appropriate dilution for 1 hour at room temperature membrane is washed with buffer containing non-ionic detergent titronics 100 and developed. Use of the 2 different antibody increases the sensitivity as well as giving flexibility to plan the multiple probing.

Then you also have 1 step probing in 1 step probing the primary antibody contains enzyme or fluorescent label for detection 1 step probing is not very common into the western blotting. Once you are done with the probing you can actually be able to develop the probe blot. So, step 6 is the blot development. So, there are multiple way to develop the blot and detect the protein present onto the membrane. So, you are going to have the different types of reagents you can actually have the chromogenic reagents or you can also have the luminescence reagents.

So, in the chromogenic reagents you can have the depending upon the enzymes for example, HRP you can actually be able to use the 4 chloro naphthol or DAB or TMB and the all of these reagents are actually going to be give you the colors. For example, 4 chloro naphthol it is the oxidized product is going to form the purple color precipitate. DAB and nickel chloride is going to form the brown color precipitate and TMB it is actually going to form the dark purple stain. Similarly if it is the enzyme is alkaline phosphatase you but you then you can use the BCIP and ABT and BCIP hydrolysis product in indigo precipitate after oxidation with NBT. If it is luminescent substrates then you can for the HRP based system you can use the luminol and hydrogen peroxide and oxidized luminol gives the blue light whereas, for the alkaline phosphatase you can

use the substituted 1, 2 dioxane phosphate and it is going to dephosphorylate substrate gives the give off the light.

So you can actually have the do 2 different types of detections you can have the colorimetric detections or you can have the chemiluminescent detections. So you wash the membrane with TBS to remove the detergents place the membrane onto the colorimetric reagent and the blot develop into 10 to 30 minutes stop the reaction by washing into the distill buffer air dry the membrane and photographs for permanent records. Then for the chemiluminescent detections the detection the different chemiluminescent reagents are given in the table transfer the membrane onto the chemiluminescent reagents soak the membrane for 30 seconds to 5 minutes strain of the reagent and wrap the membrane into the plastic wrap place it into a film cassette and expose the membrane to film to few seconds to the several hours. Then you can also have the fluorescent detection so secondary antibody labeled with the fluorescent dye and captured into the scanner can be also done. So these are the some of the method the steps what you are supposed to do when you are doing the western blotting.

We are supposed to first run the SDS page followed by the transfer followed by the treatment with primary antibody followed by the treatment with the secondary antibody and then you are going to do a washing step and then you are going to do the development. So these are the some of the very very crucial steps and you also require a lot of precautions while you are performing these steps. So I would like to take you to my laboratory where the students are going to perform all these experiments all these steps and they are also going to discuss about the different types of procedures. So they are actually going to develop the blot with the help of the chemiluminescent reagents so that you will be familiar with the whole process and so on. In this video we will demonstrate you how to do a western blot and how to analyze the result using ACL electro chemiluminescent substrate.

So here what we will do we have to run gel first then we will transfer the transfer method how to do the transfer we will show in this video. In previous video we have already shown that how to prepare a SDS page and how to run protein samples. In this video particularly we are interested in factors associated with the western blotting. For doing western blot we need membrane and transfer buffer and the transfer medium. This one is we use to transfer this gel to membrane.

So here membrane can be two kind one is nitrocellulose which has low protein binding efficiency and hydro-pulching nature. Another membrane is PVDF this is hydrophobic membrane and higher protein binding capacity. So the processing for western blot is different for nitrocellulose and PVDF. So if you are using PVDF membrane we have to

take we have to cut the part whether if you have readymade pre-cut blots then no need. If you have if you are taking from a bundle you have to cut precisely how many wells you want.

So after that you have to label front on the blot where the front side can be used for transferring the protein and that can be used in previous step further steps also like antibody incubation. So here for if you want to use PVDF membrane you have to charge with the methanol. So since the PVDF is hydrophobic membrane you cannot directly transfer the transfer in the aqueous medium. First you have to keep in methanol for at least 20 minutes. So after this can be called as charging.

So after this we will use that for transfer. So this is pre-soaked in methanol and equilibrated in transfer buffer. So here while doing transfer we need to consider few things. The buffer always should be in chilled condition otherwise during this transfer at high voltage it will generate high temperatures. So that may degrade your protein or decrease the efficiency of the transfer.

That is why we need to keep the buffer always in chilled condition and let us start the procedure. So we need a pre-run gel. So we already finished the gel running. In addition to that we also need sparges which will give cushion to the gel so that gel may not be destroyed during the transfer. So this is the cassette we will use for the transfer.

So this is negative side of cassette and this is the positive side. So we are going to keep gel on negative side and positive side the blood membrane. So when we apply voltage from this side to this side the negative protein it will be transferred it will be moved to positive side and it will be captured in the membrane. So first for doing that these sponges we need to keep and also this maybe give some non-specific binding to membrane. So what we will do we will put blotting sheets on top of this.

So after this you have to remove air bubbles if any present. So once you inserted the blotting sheet then you have to keep your gel. So here we have to remember that gel after finishing the SDS phase running you have to keep in transfer buffer so that it will give identical condition for equilibration kind of thing during transfer so that protein transfer may be easy. So this is the gel and keeping on the negative side. So after that we have to overlay with the membrane.

Next we have to remove any air bubbles if present. We have to overlay with another blotting sheet and remove the air bubbles. Each and every time when you introduce something you have to remove air bubbles. So this is the final sheet. So this is the positive side of the cassette just have to like this.

These are the screws we have to tighten it up then only the contact between the gel and membrane will be sufficient to get transferred. First you do not tighten it initially you just keep and after that press the positive side of the cassette then tighten the screws. So all these things should be done in the transfer buffer only unless specified. So this is the chilido transfer buffer.

Now we are going to do transfer. Pour sufficient buffer. Keep this ice pack also if the chilling is not sufficient then there may be heat generation. So in order to prevent that we will use this ice pack. So this will keep the buffer cool till the transfer end of the transfer. So once that is over you directly take out the cassette and keep.

If there is a buffer insufficiency you can add on top of that. Make sure that the cassette completely submerged so that the transfer will be proper and there is no air bubbles. So once the setup is over now you can transfer. Now transfer is going on. So how much voltage we need to give it depends on transfer to transfer it varies.

Generally in our lab we will give at least 2 hours of transfer at 120 volts which is sufficient to transfer even low molecular weight proteins also. But from instrument to instrument also it varies. You needed to optimize before doing transfer. After 2 hours we have to remove the blood and incubate with the blocking buffer.

So I am going to stop here. Remove the cassette. Keep the net ray. Remove the screws properly. Gently remove the sponges. Take out the blood and keep it in blocking buffer. In this condition we have to keep if you are keeping it room temperature it is for 2 hours at least.

If you are keeping in 4 degree Celsius you can keep overnight. The blocking buffer contains skim milk or BSA along with the TUNT. The next one is the Western Blot transfer. It all depends on the efficiency how precisely you are doing the transfer. For example, you should not use your bare hands while handling the blot or gel.

So whatever the proteins present on your fingers it will transfer into gel or membrane which will give high background during development of the blot. So always use gloves. Apart from that while handling the instrument make sure that there may be possibility of electricity, the shock may happen sometime. So we have to that time also we need to use gloves. And after finishing the transfer you have to clean all the apparatus properly and dry it for the next 10 years.

After the blocking of the membrane we have to remove the membrane and incubate with

the primary antibody without washing. The main purpose of the blocking is that it will occupy non-specific sites other than the respective protein so that when the antibody comes it will bind to that specific protein and gives no noise. So after this we will incubate with the primary antibody for overnight at 4 degree Celsius then wash 3 times at least 15 minutes each with the TBST buffer or PBST buffer and again treat with incubate with the suitable secondary antibody. For 5 hours at 4 degree Celsius or 2 hours at 3 hours at room temperature. After that we need to wash properly at least 3 times then we will develop the blot with the electrochemiluminescent substrate.

In earlier western blot how to do western blot we showed how to transfer the proteins to membrane. So we incubated with the primary antibody following secondary antibody and wash with the. Now here we show how to develop a blot. For developing a blot we need chemiluminescent substrate. In most of the commercially available kits Luminal is the one of the substrate we used for this purpose.

So Luminal in presence of hydrogen peroxide and peroxidase agent which present in the secondary antibody. This peroxidase conjugated secondary antibody. This peroxidase converts Luminal to excited Luminal by deprotonating and oxidizing it. So this excited product gradually leaves the energy by releasing luminescent products that light will be detected using this instrument. So these are the commercially available chemiluminescent substrate solutions.

So it is available from wide range of companies. We have to mix 1 is to 1 ratio. So we have to take out the blot, drain the buffer whatever present properly. So after that we keep blot in between plastic paper foils. Then we will take chemiluminescent substrate.

So after that we have to slowly press and remove air bubbles. This is the tray we used for developing the blot. So we have to open the system, properly align the tray and then shift blot to the table. Once it is over, we have to just close. Here we have to select application. We want blots that is chemiluminescent and what exposure want? You have two options, manual auto.

In auto two options are there, optimal auto exposure, rapid auto exposure. We will choose optimal auto exposure. So you can enlarge the blot also.

Once it is over, you just say. So this is the developed blot. So as we can see the bands pattern. So this is how we develop western blot through electro chemiluminescent substrate. So in this video we have demonstrated how to transfer proteins to a blot and what are the precautions need to be taken while doing the western blot and also how to develop the blot and what is the laying principle behind the developing the blot. So I

hope this will help you to understand the basic outline mechanism of how western blot works. So what these students have discussed? They have discussed about all these steps.

They have discussed about how to transfer the proteins onto the membrane, how to perform the primary antibodies, how to perform the washing so that and what are the different precautions you should take while you are performing the western blotting and so on. So these are the different steps what you are supposed to perform. Now so what we have discussed so far? We have discussed about the southern blotting. We have discussed about the northern blotting.

We discussed about the western blotting. The purpose of these technique is very clear right for the southern blotting it is actually going to detect the DNA. So what if you see if you try to correlate how you are actually going to answer the different questions with the help of the blotting techniques. So the southern blotting is actually going to tell you that whether a particular gene is present or not right. So it is actually going to tell you about the gene is present or not.

But whether this gene is actually expressing or not right. So whether this gene is performing the transcription or not that information you are going to get with the help of the northern blotting. So northern blotting is actually going to tell you whether this gene is forming the RNA or not right. Because there are genes which are also present but they may not be expressive genes they may be only present into the genome simply by no other reason or they may be expressing but in that condition in which they may not be high skipping genes they may be required for a specific purpose so they may not be expressing. So northern blotting is actually going to tell you whether the RNA is forming or not. So it is actually going to tell you about the transcriptional activity of that particular gene.

So it is actually going to tell you the transcriptional activity of the gene. Then whether the gene is forming the protein or not right. Because there are proteins there are genes which are not expressing there are genes which are expressing but at the only at the RNA level they are not translating into the protein. So that information you will get by the western blotting. So what you see is so western blotting is going to tell you whether there is a translational activity of that translational activity of gene is present or not. So what you see that with the help of these three individual blotting techniques you can be able to get the complete information about a particular gene.

You can actually be able to know how many number time this gene is appeared into the genome right whether the gene is present in genome or not number one number two what

is the location of this particular gene and how many times this gene is present into the genome. Then second question comes that whether this gene is forming the RNA or not whether gene is transcriptionally active or not and that information you will get from the northern blotting. And then whether the RNA what it is forming is expressing and being utilized by the protein synthesis machinery to give you the protein that also can be answered with the help of the western blotting. So all these western all these blotting techniques are complementary to each other they are actually going to give you the in-depth knowledge about that particular gene fragments and so on. So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss some more molecular biology techniques. So till then we are going to conclude our lecture here. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 09
Molecular Techniques (Part 2)
Lecture-37 Polymerase Chain reaction (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati. And what we were discussing, we were discussing about the different aspects of the molecules into the course molecular biology. So, for what we have discussed, we have discussed about the cell biology, we have discussed about the cellular structures organelles their role into the cellular metabolisms and then we also discussed about the different types of biomolecules. So, we have discussed about the proteins, lipids, enzymes, DNA and RNA and we have also discussed about the their structure and functions and how they are regulating the different events. And then subsequent to that we have also seen the how these molecules are even regulating the cellular events.

So, we have discussed about the apoptosis, we have discussed about the autophagy and so on. And we also discussed about the cell division and then subsequent to that we have also discussed about the central dogma of molecular biology and we have discussed in detail about the different events what is comprising the central dogma of molecular biology. So, we discussed about the replication, transcription and translations and following that we have also discussed about the different types of techniques. So, when we were talking about the molecular techniques in the previous module, we have discussed about the blotting techniques and purpose of the blotting technique is to detect the particular biomolecules.

So, in the southern blotting we are going to detect the presence of a particular gene fragment into a particular DNA molecule or the genomic DNA. Then we discussed about the northern blotting. So, the northern blotting is going to tell you about the presence of that particular gene product. So, we are going to talk about the transcriptional activity of that particular gene and then we also discuss about the western blotting. So, western blotting is going to talk about the expression of that particular gene.

So, in the previous module we discuss about the blotting techniques. In the current module we are going to discuss about the some of the amplification techniques. So, in the amplification technique one of the most popular amplification technique the polymerase chain reactions. So, what is the polymerase chain reaction? So, polymerase chain reaction is a technique which actually allows you to the production of or the synthesis of a specific sequence from the genomic DNA or the other source of the DNA. Now, before we get into the detail of how to this technique is been developed, how this technique is been evolved, it is important that you should actually be able to understand the process of DNA synthesis.

And for that I would highly recommend that if you have not seen the my previous video of the DNA applications then you should actually should go through because the sole basis of the PCR

is the fundamental steps what is been found when the people have discovered the DNA applications. And more or less the machinery also remains the same except that some of the events are been modified where the protein was required. So, it has been modified with the physical parameters so that you can be able to achieve the similar results. So, it is important that you should be able to understand first the DNA application in detail so that in fact, you should also read about the DNA structures and other things. So, that it will be easy for you to understand the process.

Now, what is PCR or polymerase chain reactions? So, polymerase chain reaction is been used to amplify a double standard DNA molecule from the genomic DNA and it is been required to amplify a double standard DNA with the same size and sequence by the enzymatic method and the cyclic method. So, for example, you can imagine that if this is the original DNA right what is mean by the PCR is that this is the original copy of the DNA you are actually going to generate the multiple copies of this particular DNA. So, for example, I have generated first copy, second copy, third copy, fourth copy all these four copies are of having the same sequence what is present in the original copy. So, this is actually the original DNA what I have used in this case and I want to amplify this DNA. So, what I have done is I have amplified this DNA and I have made the four copies now and with the same size and the same identical sequence and this is been achieved by the process of called as polymerase chain reactions.

Now, when you want to understand the polymerase chain reaction it is as I said you know it is important to study the DNA structures and the DNA applications. So, the process of this polymerase chain reactions or the basic understanding why you can be able to do to you why you how you can be able to do this is actually lies into the DNA structures. So, do you know that the DNA is actually the complementary in nature. So, you are going to have the primary strands. So, this is the first strand and then you also going to have the complementary strand.

What is mean by the complementary strand is that whatever you are actually going to be present here it is mean going to be complementary and complementary is going to follow the rule that A is actually going to have the two base pair or two hydrogen bonding with T and G is actually going to have the three triple bonding with the C. So, wherever you have A into the primary strand it is going to be T in the complementary strand. Similarly, wherever you have G in the primary strand it is going to be the C in the complementary strand. This means if for example, you have G here. So, you have C here, you have G here, you have C here, you have T here.

So, if you have T here in the primary strand then it is going to be A into the complementary strand and so on. So, if you have A in the primary strand then it is going to be T in the complementary strand. So, if I want to amplify this particular sequence what I can do is I can use this DNA replication machinery and I can be able to use this as a template for the leading strand. So, for example, this is going to be a leading strand and then this is going to be a lagging strand. So, I can actually be able to do the synthesis of this into utilizing this as a template and I can use the synthesis of this utilizing this as a template.

And that is how I can actually divide this into two parts. I can divide this into two sequence. So,

you can have this is the 5 prime to 3 prime and the other one is 3 prime to 5 prime. So, this is the first this is the complementary strand, this is the primary strand what you are going to see here. And then what I can do is I can just use this and I can just put the DNA replication machinery and it is actually going to give me this particular sequence.

Similarly, I can put the DNA replication machinery on this side and I can actually be getting this. So, if I if you recall the DNA replications so that DNA replication machinery only can be used for achieving this particular target. So, that you can be able to have the same sequence identical sequence been present as multiple times. So, this can be done by multiple times and that is how you are going to have the amplified DNA. So, for this you actually need to run the DNA replications.

Now, if you see the DNA replication, DNA replication has three steps. You have initiation, you have elongations, you have terminations. So, in the initiation the DNA is going to be you know prepare. So, you are going to prepare the template DNA if I write like this. So, if you are going to have prepare the template DNA, that template DNA is preparation involves the protein binding to the DNA and then you are going to have the opening of the double helix, which means at this stage you are actually going to open and this you know break the hydrogen bonding between the binds.

And that is how this is going to be 3 prime to sorry 5 prime to 3 prime strand and this is going to be 3 prime to 5 prime strand. So, both of these strands are now going to be prepared for the synthesis. Now, in the elongation step you are going to have the synthesis step. So, in the initiation step also remember that the DNA a primer is going to be synthesized. So, in this case also DNA primer RNA primer is going to be synthesized and then this RNA primer going to be used by the DNA replication machinery and that is how DNA replication machinery is going to sit and it is going to synthesize the DNA as per the complementary T from this particular DNA right.

Similarly, the DNA strand will synthesis will stand here and it is going to synthesize this one right. So, in the second stage you are going to have the elongation which is going to be synthesis phase. So, the protein connects the correct sequence of the nucleotide into a continuous new strand of the DNA and then you are going to have the termination. So, it is going to stop the DNA synthesis. So, protein release the DNA application is over then the it is actually going to end up into the termination and that is how it is going to stop the DNA synthesis and protein are going to be released from the replication complex.

Now, for these three events you requires a battery of the enzymatic factors which are present in the cell and they are present in the correct ratio. So, that they one or other factor is going to come and this these events are actually going to be completely the controlled. So, let us see what are the different protein machinery what is required for initiation, elongation and terminations. So, you require the helicases, you require the primase and you require the single standard DNA binding proteins this is for the initiation ok. So, this is going to be the for first event that is the initiation then you require the DNA application and that is required for the elongation and you also require the tethering protein that is also required for elongation and termination is

always been done by the TIRP proteins.

So, helicases it is going to separate the two strands right that is how it is actually going to melt the DNA and melt DNA is going to have the two single standard DNAs and then the primase is actually going to have the RNA primer synthesis that so that it is actually going to provide the attachment point on to the DNA and that is how this attachment point is going to be utilized by the DNA polymerase. Then SSB, SSB is going to prevent the re-endling of the single standard DNA or strands then you are going to have the DNA replications or DNA polymerase which is actually going to be participate into the synthesis of the new strands and then you also require the some of the tethering proteins that is actually going to stabilize the polymerase on to the DNA. Now, remember that all these enzymatic factors are present inside the cell and these factors are being present in a very adequate amount and there is a complete you know cellular machinery which is taking which is actually governing the presence of helicases, presence of primase, presence of SSB proteins. So, if I want to replicate these events into in vitro synthetic system then I need to add the helicase, primase, SSB at a regular time intervals right at a regular time interval and all these proteins will have their own half-life. So, it is important that it is very difficult to even manage these the regulation part right.

For example, when the helicase is separating the two strand the primase and SSB are actually going to be keep standing they are not doing anything once the DNA is been single-stranded and once it is going to open then only the primase will come and sit and it will actually going to synthesize the RNA primer right same is true here and this also right and while it is doing so SSB is actually going to go and bind the single-stranded DNA SSB will go will cover the nucleotides right that is how the it happens when you are doing the DNA applications and these SSB molecules are keep detaching from the DNA strands while the this been taken care by the DNA polymerase right. So, when the DNA polymerase when the primase will go and the primer is already been synthesized the DNA polymerase will come and sit and do and when the DNA polymerase is running it is actually removing all these SSBs right. This means SSBs are now been present into the solution right. Now helicase is also present in the solution right. So, helicase will actually going to interfere and then we do not know what will be the stoichiometry of these machinery which is required for DNA applications and so on.

So, that is why it is important that some of these events can be modulated by a physical parameters. So, that is what you are supposed to you know achieve when you want to do the polymerase chain reactions under the or the DNA application under the in vitro syndrome. So, if you want to achieve that you are actually going to do that by simply achieving all these three events in a in a by another method right. So, for example, for initiation you required a helicase in the natural system right, but the same can be achieved simply by heating the DNA because what helicase is doing it is actually breaking the hydrogen bonding right it breaking the hydrogen bonding between nucleotides. Same can be achieved by heating the DNA and you if you heat the DNA and if you heat the DNA at 95 degree Celsius or 97 degree Celsius then what will happen is that heating is also going to achieve the same thing it is actually going to break the hydrogen bonding between the nucleotides.

This means you can easily bypass the role of helicase simply by heating the DNA at 95 degree Celsius that does not mean that in the cellular system also that can be achieved. So, in the cellular system everything is being done at room temperature or 37 degree Celsius and that is why you require an enzyme to perform this task, but in the case of in vitro system because DNA is fairly stable right. So, you can actually heat the DNA and once you heat the DNA strands are going to be get separated and they will remain separated until the temperature is high and then you can enter into the another phase which is called as elongation phase and in the elongation phase you can actually be able to cool down this temperature right you can cool down to such extent the DNA will still remain as single standard, but it will not going to anneal it is not going to reanneal to each other and give you the double standard DNA and then what you are going to do is you are going to add the DNA polymerase you are going to add the DNA polymerase right and DNA polymerase is going to start the synthesis by taking the primer into the consideration and it is going to start the synthesis right. Once the synthesis is over and you can calculate how long this will take for doing the synthesis till this end you can actually be able to again go back to this and you can increase the temperature and this can continue for some cycle and that is how you can actually enter into the third stage where you can actually cool down the temperature and this can be achieved by keeping the reaction at room temperature or at 4 degree Celsius. So, once you bring the temperature at room temperature and 4 degree you are ideally going to do the termination.

So, that what you see here is that the DNA synthesis under the in vitro condition can be achieved and you can be able to remove some of the crucial enzymatic factors what is been participating very extensively into the DNA applications, but you can actually be able to replace them by simply changing the physical parameter and that is actually was a groundbreaking discovery because that has discovered a technique which is called as the PCR or polymerase chain reaction. But as I said you know these any technique what you are actually going to devote or evolve does not come into a single day it is actually going to come into a very very long day of doing the different types of discoveries. So, let us see how the technique is been evolved over a course of you know couple of decades. So, there are different events which are been responsible for the development of PCR. So, in the case of 1950 there was discovery of mechanism of DNA replication by the Arthur Kornberg.

He discovered the first DNA polymerase and other factors like helicases and primases. So, it is started with the DNA replication it started with the basic understanding about the DNA replication and then the people have discovered the machinery right just now that what we have discussed right you require the helicases, primases, SSBs, DNA polymerase and so on. So, all that machinery is been discovered in the early stage of 1950 right. And then what you see here is there is a huge gap of so in the year of 1976 the people have discovered the first thermostable DNA polymerase and you will understand this when I am going to discuss about different steps of the polymerase chain reaction why it is important to have a thermostable DNA polymerase. Then in the year of 1983, Kanimoulas and synthesizes the DNA oligos probe for the sickle cell anemia mutations and that is how he has performed the polymerase chain reaction to amplify the particular gene what is responsible for sickle cell anemia and that is how he could be able to show or demonstrate that you can be able to perform the DNA applications or you can be able

to perform the amplification under the in vitro conditions.

And in the same year 1983 he also discovered how you can be able to do the repeated thermal cycling and so that you can be able to know amplify a small segment of the cloned gene and you can be able to use that for reducing the large amount of DNA. Then in the year of 1984 the Kanimoulas and Tom White tried a different experiment to test the PCR on the genomic DNA and but the amplified product was not visible into the agarose gel. And in the year of 1985 the patent was filed for PCR and its application focusing on the sickle cell anemia mutations. And in the year of 1985 the use of thermostable DNA polymerase in PCR was started out of only two enzymes Taq and BST known at the time. The Taq was more found more suitable for the PCR.

And then in the year of 1985 the first announcement of PCR technique in Salt Lake City and then in the year of 1985 to 87 there were development of instrument for the PCR and its reagents. These are some of the notable discoveries I am not saying that I am going to give you a comprehensive events or comprehensive account of the events what is being discovered or what is being related to the PCR. Because there are many more probably things are being done apart from these things. But these are some of the important discoveries what has been done and that has been responsible for the development of the polymerase chain reactions.

Now let us come to the technique. What you are going to do is you are actually going to do a repeated cycling of the same replication. And that is how you are actually going to do the amplifier. How you are going to do that? You are going to do that by using a you know by using the simple technique like for example, this is the this is actually the template DNA ok or this is the starting DNA I will say template or starting DNA. Now this is a double standard DNA what you are going to do is you are going to do a denaturation. So, once you do the denaturation it is going to denature as a single standard DNA and it is going to form the two DNA strands.

This is going to be the strand 1 and this is going to be the strand 2. So, for example, if this is from 5 prime to 3 prime and this is the 3 prime to 5 prime then you are going to have this as 5 prime to 3 prime and this is the 3 prime to 5 prime. Now what you are going to do is you are going to add the primers right. So, you are going to add the primer. So, what will happen is the primer will go and sit on to the 3 prime end right.

So, it is going to sit on to 3 prime end. So, this is going to be on 3 prime end. So, PC has how this is going to be sit here right and then you are going to change the condition. So, that it will enter into the elongation phase you are going to add the DNA polymerase. So, it is going to utilize this primer and it is going to synthesize and it is going to go up to what because the template is going to be over. So, it is going to synthesize up to this same is true for this one also it is going to synthesize up to this one right.

Now what you are going to get is you are going to get the two copies. So, you started with one copy right and now you are after the first cycle. So, this is going to be after the one cycle you are going to get 2 DNA molecule right you are going to get 2 DNA molecule this is the one molecule and this is the second molecule. Now again if you continue the same thing right it is actually.

So, what you see here is that you are going. So, you started with the one DNA molecule you are going to have the 2 DNA molecule at after the end of the first cycle. Now if I do the same event again and again in the second event what will happen is this is actually going to give me the two templates right. So, it is going to give me the two templates imagine that this is same way I am going to show you right and again the same thing will repeat same is true for this one also this is also going to give me the two template this means after the end of the second cycle I am going to get the 4 DNA molecule and after the end of the third cycle I am going to get the 8 DNA molecule right and this will continue like after this you are going to get 16 molecules after 16 you are going to get 32 molecule and so on. So, after every cycle it is actually going to give you the double amount of DNA because the previous template previous synthesized DNA is also going to serve as the template right.

Now this is what is actually been achieved. So, PCR is a repeated cycle reaction that involves the mechanism of DNA replication. It results in the product of multiple copies of a DNA from a single one. The whole process involves three main events denaturation, annealing and elongation. A DNA fragment of interest is used as a template from which a pair of primer or a short oligonucleotide complementary to both the double standard of the DNA are made to prime the DNA synthesis where the direction of the synthesis or the extension is from 5 prime to 3 prime this means it is actually going to start from the 3 prime end as a template right as in as in in a DNA applications. The number of amplified DNA or the amplicon increases exponentially per figure such as the molecule of DNA give rise to 2, 4, 8, 16 and so forth after every cycle and if you want to calculate the amount of DNA right what you are going to amplify it is going to give you a with this equation which is called C is equal to $C_0 e^{nt}$ where C is the final amount of DNA C_0 is the initial amount of DNA and E is actually the efficiency and n is the number of cycle and S is the slope of the exponential phase.

So, E is always been calculated as 10^S and if E is 1 then the S is going to be 3.3 to 1. So, this is actually going to be used for calculating the amount of DNA what you are going to get because remember that after every cycle it is actually going to get amplified at double right. So, it is from 16 it is going to be 32 then 64 and so on. So, it is going to be an exponential application and remember that it is going to have the three events denaturation, annealing and elongations.

So, let us discuss about these events. So, that you will be able to understand more better about the DNA polymerase or polymerase chain reactions. So, polymerase chain reactions is going to be have the denaturation, annealing and elongations right. So, it is going to have the two round of denaturations you are going to have the initial denaturation which is actually going to denature the whole DNA. So, example you are talking if you are doing this from genome. So, then it is actually going to denature the whole genome right.

In the further denaturation it is actually going to only denature the small amount of fragment or a small fragment from where it is the primary is actually going to bind. So, in the initial denaturation it is going to be very long. So, you can heat the PCR mixture at 94 degree Celsius to 96 degree

Celsius for 10 minutes to ensure the complete denaturation of the template DNA. So, it is going to denature the complete genome and you are going to have the single standard two copies of the genome right. And then you are going to have the these events which are called denaturations, annealing and elongation on a repeated mode right.

So, these are going to be a part of the cycle right. So, you can see this is the initial denaturation right which is actually going to be called as stage 1. So, for example, in this case I have we have done it 95 degree Celsius for 5 minutes. So, it can be 5 minutes it could be 10 minutes. Then in the stage 2 you are going to do a repeated cycling. So, this is going to be a cycle this is a one cycle where you are going to and you are running this cycle for 30 times right.

In a single cycle you are going to have the denaturation, annealing and elongation. So, what happened in this denaturation? So, this is the first step in which double standard DNA template is denatured to form the two single standard DNA by heating at 95 degree Celsius for 15 to 30 seconds. So, for example, in this case it is 95 degree Celsius for 30 seconds. Then you are going to have the annealing. So, this is the annealing step where the temperature is going to be lower down not to that extent that the single standard DNA is going to be double standard DNA and going to form the double standard DNA, but good enough so that the primers and the should be able to anneal or should be able to bind the template according to the sequence complementary.

So, you are going to have usually 50 to 65 degree Celsius primers are allowed to bind the template DNA annealing time is 15 to 30 seconds. You remember that this time is being regulated in such a way so that it should actually allow the annealing of the primers to the template, but it should not allow the complementarity DNA should be you know should not form the double standard DNA. And then it depends on the length and the basis of the primers. And then you are going to enter into the elongation force. So, this is the synthesis step where the polymerase perform the synthesis of the new strand in the 5 prime to 3 prime direction using the primers and the dNTP.

So, dNTPs are the raw material what you require for the synthesis and average DNA polymerase adds about 1000 base pair per minute and step 1, 2, 3 make the 1 cycle and in general 30 to 35 to 40 cycles are performed in a typical PCR amplification. So, this is actually denaturation annealing elongation is going to be performed on a repeated mode. And for example, in this case, we are doing it for 30 times, but you can do somewhere around 35 to 40 cycles. After the cycles are complete, which are all the 30 cycles are complete, then you are going to have the final extensions or the extra elongations. And the reaction is held at 70 to 74 degrees Celsius for several minutes to allow the final extension of the remaining DNA to fully extended and then you are going to do the termination.

So, finally, you are going to hold the reactions at 4 degrees Celsius. So, that is actually going to terminate the synthesis and it is going to keep the reactions at 4 degree for a very very long extended period of time. Now, if I want to perform the PCR reactions, I first require to perform up you know, the setting up these cycles and the setting of these cycle has to be done into a PCR machine. So, what are the these cycles? So, you are going to have the initial denaturation which

is 95 degrees Celsius for 5 minutes, then you are going to have the you are going to set up how many number of cycle you want to perform. So, for example, if I want to perform the number of cycle at it for 30 cycles, then you in 30 cycles, you are going to have the denaturation, annealing and extension.

So, you are going to have the annealing for 95 degrees Celsius for 30 seconds, 55 degrees Celsius for 30 seconds and extension would be 72 degrees Celsius for 45 seconds. All these annealing and extension depends on the annealing extension especially if the extension depends on the length of the gene or length of the gene fragment which you are going to synthesize. So, as I said you know 1000 base pair per minute that is the criteria. So, if you are synthesizing for example, 1000 base pair right, then it is going to be 1 minute. Then you are going to have the final extension that is the 72 cycles for 5 minutes and then you are going to have determination or the holding that is usually been done at 4 degrees Celsius.

Now, what is the requirement or what are the different things you require for performing the PCR? So, if you want to perform the PCR, you actually require following information or following materials. You first require the DNA sequence of the target region must be known. So, that is very very important because and that is probably the one of the drawback of the PCR that you require the sequence of this particular DNA what you want to amplify. Then you require the primers. So, typically 20 to 30 base pair in size they can be readily produced by the commercial companies can also be prepared using a DNA synthesizer.

So, then you require the primers. These primers are been always been directed against this particular sequence and that is why you require this particular sequence. Then you require the thermostable DNA polymerase for example, the tag DNA polymerase which is not inactivated by heating at 95 degree Celsius. So, remember that when we were discussing about this right. So, if you do not have a thermostable enzyme or thermostable DNA polymerase then the moment you are going to enter into the denaturation state that is the 95 degree Celsius the DNA what the enzyme what you have added like the DNA polymerase is going to be inactivated. So, to avoid this only people have discovered the thermostable DNA polymerase.

So, that it when you are going to run these cycles for 30 cycle or 40 cycle when you reach to the denaturation stage right then you want to then your enzyme should remain intact and enzyme should not be get inactivated because then you are supposed to add the enzyme in a repeated cycle and then it is very very impossible or it is difficult to perform the PCR reactions. And then you also require the DNA thermal cyclers. So, machine which can be programmed to carry out the heating and cooling of the sample over a number of cycles. These thermal cyclers are actually very specialized machine and their job is nothing but to regulate the required temperature. So, that if they should incubate your reaction at a particular temperature until the people who have not discovered the thermal cycler what they were doing is they were you know manually changing the reactions they were manually changing the tubes from one temperature to another temperature.

For example, in the initial stages of PCR development they were having the three different

buckets right. So, they were having the three different types of incubators or water baths one is set as 95 degree Celsius the other one is as set as 55 degree Celsius and the third one is set at 72 degree Celsius. This one is actually doing the denaturation this one is doing the annealing and this one is doing the elongation. So, what they were doing is they were setting up the reaction into a you know a block right. And then this block is either first placed into the 95 degree Celsius then from here you are they are going to place it into here and then from here it is going to be place it here right.

And then from here again it will come back to this place again like that. So, it was very very cumbersome it was very difficult to perform and then during this you know shifting of the block from 95 degree Celsius to 55 degree Celsius or from 55 to 72 there will always been a change in temperature and other kinds of things happens and that also affects the overall productivity and overall quality of DNA what they get actually. But that time the electronics was not intact and we were not having the very good quality you know cooling machines and all that and that is why the people are not having the this kind of sophisticated thermal cyclers. So, this is a typical thermal cycler what you see is this is from applied bio system right. And the thermal cycler is a instrument that carry out the amplification via the polymerase chain reactions.

The device has a thermal block. So, what you see here is this is a thermal block and that has the holes right that has the wells. These wells you are actually going to place the your appendops or the reaction vessels. So, device has a thermal block which has holes where the tubes holding the reaction mixture can be inserted. The cyclers then raises and lower the temperature of the block in a discrete pre-programmed steps and you have seen the programming right that you are actually going to ask the system to you know increase the temperature to 95 you lower down the temperature to 55 and then hold the temperature at 55 for 45 seconds 30 seconds or whatever. And that can be done very efficiently into this kind of program in this kind of machines.

Now, how this machine actually do? So, these machines are always been using the Peltier system and the Peltier system is cooling and heating the system or very efficiently. It is much better than the compressors actually. So, each module is composed of the N and P type of thermal semiconductor material and the Peltier phenomena occur at the junction of these two similar conductors. So, when the current is passed to the junction a transfer of heat occurs from one side of the junction to the other and QL is the absorb heat on the cold side and the QH is the release heat on the other side. So, because of this the you know the absorption of heat or the release of heat you can be able to maintain the temperature very precisely and that too at a very rapid rate.

Now, let us move on to the what are the material required for the performing the polymerase chain reactions. So, as I said you know you require a template DNA. So, you can have the 1 picogram to 1 nanograms of template DNA for a viral or the short templates, but you require a 1 nanogram to 1 microgram for genomic DNA. So, depending upon the source of the DNA you can be able to use the different types of products, different amount of DNA actually. So, you can use 1 picogram to 1 nanogram if it is a viral or the short DNA short templates, but if it is a genomic DNA then you are supposed to use 1 nanogram to 1 microgram.

Then you require the primers. So, you require two different types of primers you require a forward primer and a reverse primer which means you require a primer for doing the synthesis of the leading strand and you require a primer for doing the synthesis of the lagging strands. Although there is no leading and lagging strand in the case of polymerase chain reaction because the synthesis of both the strands actually completed at the same time, but just for making you more familiar with the DNA replication I am using this kind of terms. So, you require two different types of primers forward primer and a reverse primer and you are going to add the primer in the range of 0.

1 to 0.5 micro moles. Then you also require the magnesium chloride because magnesium chloride is going to work as a cofactor for the DNA polymerase. So, you require the 1.5 to 2 millimolar DNA polymerase for the Taq DNA polymerase and then you require the dNTPs. So, you require the four different types of dNTPs you require the dATP, you require the dTTP, you require the dGTP and then you also require the dCTP. So, you require the ATGC that is what is required for the synthesis of the DNA and the typical concentration is 200 micro mole of each dNTPs.

Then you also require the Taq DNA polymerase, so Taq DNA polymerase you require the 0.5 to 2 units per 50 micro reactions. Remember that the PCR is going to be performed into a large volume and I will explain you this when we were going to talk about the restriction digesters and other things. So, remember that when we were talking about restriction digestion I said you know that restriction digestion has to be performed into a larger volume such as 50 microliters whereas the PCR you are supposed to perform also into the larger reactions, but ligation when you are going to do it has to be done in a smaller volumes. Now, let us talk about the template DNA. So, template DNA could be of multiple types it could be a genomic DNA, it could be a DNA fragments, it could be plasmids, it could be a viral DNA or it could be a tissue sample.

Whatever would be the source you are supposed to use the amount of DNA accordingly it should be a blood actually, it can be a blood sample, it could be a blood sample from the crime side or some place. So, that also can be used. So, from the blood sample also you are going to use the DNA for detecting a particular fragment. So, these are the different types of sources of the DNA templates. You can have many more things from the tissue you are actually going to get the cells and from the cells you are supposed to isolate the DNA and that is going to be served as a template for performing the preliminary chain reactions.

Then you require the primers. A primer is a short DNA stretch that serve as a starting point for the DNA synthesis. In PCR you require the two primers which bind each of the single standard DNA flanking the target sequence. These are called forward primers and the reverse primers. The primer have a sequence complementary to the sequence in the template strand where they are supposed to start the synthesis.

Then you require the enzyme. So, you require the Taq DNA polymerase. So, Taq stands for the thermos ecus which is a microbe found in the 170 degree Fahrenheit hot springs in the

Yellowstone National Forest and the Taq DNA polymerase is stable in a high temperature and act in the presence of magnesium. The optimal temperature for the Taq DNA polymerase is 70 degree Celsius. Now, what is the December? These are the advantages that it will not get denatured when you are doing the different cycles and all that, but it also has the disadvantage. So, when you are actually going to use the Taq DNA polymerase, it is going to have the different types of determinates. Taq polymerase lacks 3 prime to 5 prime exonuclease proofreading activity commonly present in the other polymerases.

So, it means it is actually not going to check whether I have added the correct nucleotide or not. So, it is actually going to make you, it is going to synthesize the DNA, but may have the errors. Then Taq DNA polymerase misincorporate one base pair into the 10 to power 4 base pairs, right, which means out of 10,000 bases, it is going to incorporate one base pair. Although this number looks small, but when you are going to do a synthesis of the DNA synthesis of a large DNA fragments, it could actually accumulate the mutations, right. In the first cycle, it may mutate the number 5, it may in the second cycle, it may mutate the number 8, it may third cycle, because you know, this is, this we are talking about in the each cycle.

So, for example, in the first cycle, it may mutate the fifth residue, right. In the second cycle, it may mutate the 8th, right, it may mutate 11th. So, that is how it is actually going to keep accumulating the mutation when you are doing the 36 or 37 cycles. So, that is why it is important that we should not have these kind of mis-corporations. For a 400 base pair target with contain an error is 33 percent of the molecule with the 30 cycles. So, if you are actually amplifying a 400 base pair DNA, right target DNA, then it is actually going to incorporate 33 percent of the new molecule after.

So, whatever the molecule it is going to produce out of those the 33 percent molecules are going to have the, the you know, mis-incorporated nucleotides. So, error and the error distribution is going to be random. This means what we are going to be emphasize is that if you use the tag DNA polymerase, it is actually going to incorporate the mutations into the amplified DNA and that is not going to good if those mutations are going to change the downstream applicable, the applicable utility of that DNA for the downstream applications. So, what is the alternative? Alternative is that you are going to use the other kinds of DNA polymerases. So, you can use the Pfu DNA polymerase which is from the pyrococcus furioces and it process the 3 prime to 5 prime exonuclease proof reading activity and because of this it actually can cross verify whether it has added the correct nucleotide or not.

The error rate is only 3.5 percent after 20 cycle more amount of primer is added to avoid the primer dimerring and for an as explored gene the primer used in the closely related species are being used. Remember that what we have said, we have said that the prime requisite of a polymerase chain reaction is that you should have this information about the sequence of the DNA. But what if you do not have the information about the gene information or you do not have the sequence information, then in those cases you can be able to synthesize the primer based on the closely related species. So, if you use the closely related species for example, you do not know the sequence in human, but you may know the sequence in monkey or you may. So, that

sequence information you can use and it may actually give you the amplified product and that you can actually be able to sequence and verify whether it is actually that sequence or not and that is how you can be able to overcome this kind of limitations.

Now, let us see how you are going to set up the PCR reactions. So, the first setting up the PCR reactions you are actually required the template DNA, primers, magnesium chloride, ENTPs, tag DNA polymerase, buffers and water. So, what you are going to do is in a 50 microliter reactions what you are going to do is you are going to first going to add the water, whatever water as per the calculations, then you are going to add the buffer, tag DNA polymerase buffer, then you are going to add the DNTPs, you are going to add the primers, you are going to add the magnesium and then you are going to use the tag DNA polymerase and that is actually going to make you the PCR mix. Now, into this PCR mix you can be able to add the template. So, you can actually be able to use the different types of templates if you want and that is why you are going to add the template at the end and as soon as you add the templates then you are supposed to put it into the thermal cycler and you are going to close this lead.

This lead the top lead also has a heating block. So, this what you see here is this is actually a heating block and it is actually going to stop the evaporation of the material from the your appendops or your tube actually. So, because the top heating block is also going to be maintained as 100 degree Celsius. What is the role of this? The role of this is that suppose this is a tube and you have kept the PCR reaction at the bottom. So, this is going to be placed in this block which is actually going to maintain the temperature of 95 degree Celsius, 55 degree Celsius, 70 degree Celsius.

So, in that case what will happen is it is actually going to evaporate although it is closed by the lead. It is going to be closed by the lead, but it is actually going to evaporate. So, all these are going to evaporate and they will actually going to accumulate on the top. It is going to accumulate on the top surface. Now, if it is accumulating on the top surface then it is all the enzymes and other things will not get a chance to interact with the reaction mixtures. So, to avoid this what you are going to do is you are going to keep a heating block on the top and what is the temperature of the heating block? The temperature of the heating block is 100 degree.

So, even if you have a temperature of 95 degree Celsius at here the top temperature is 100 degree. So, because of that it will all the liquid will come down because it will evaporate, but the top chamber it the temperature is more. So, that is why it will not go up it will actually be present at the bottom only. So, it will all the liquid will remain as the liquid in the tube and that is why it is actually going to perform the reactions. Now, let us take I will take you to my lab for a demo how you are going to set up the PCR reactions, how you are going to perform the PCR reactions and how you can be able to set up the reactions and analyze the results. In this video, we will be demonstrating how to set up a PCR reaction and analyze the results using other room gel electrophoresis.

PCR or polymerase chain reaction is a widely used molecular biology technique to amplify a particular segment of DNA. It is also employed in biomedical research and forensic medicine.

The main application of this is polymerase chain reaction is ronin. To set up a PCR reaction, we need template DNA, site specific primers, DATP mix, nucleus free water and Taq polymerase. For a 50 microlitre reaction in a typical concentrations of 10 to 100 nanograms of template DNA used and 5 picomoles of each primer will be used.

So, this is an earlier version of thermal cycler which contains display unit where we can observe the parameters and change the parameters. This is a hard shield, this is sample holder and inside there is a Peltier system which can maintain the temperature fluctuations. For setting up a PCR reaction, initial denaturation at 95 degrees Celsius, 3 minutes and these steps we will use 30 repeats where initial denaturation will be 30 seconds and annealing it, extension time should be given 1 minute per kB. And here final extension should be given 10 minutes and hold at 4 degrees Celsius, 10 minutes.

Once the PCR reaction is completed, we have to analyse the results for amplification. For that we need agarose and TAE buffer. First we have to weigh agarose and mix with the TAE buffer. It will not dissolve easily, so we have to heat it in microwave oven until it get dissolved.

Now agarose got dissolved in TAE buffer. We have to let it cool down up to 53 degrees Celsius. Now before pouring we have to add lithium bromide for detection purpose. Now the gel got solidified, we have to take out the gel and keep it in the electroporetic apparatus. We have to gently remove the comb, lose the knobs and keep the gel in the apparatus.

Make sure that the buffer is submerged in the gel. We have to fill the remaining part with 1X TAE buffer. Generally for analysing the DNA samples we will use agarose gel electrophoresis. This is the power pack and this is the electrophoretic apparatus. This is negative electrode and this is a positive electrode.

We can change the voltage from here. For loading of sample, we have to mix PCR reaction mixture with 5X loading time. So after loading is over we have to cover the electroporetic apparatus with the lid and we have to adjust the voltage. Then start rough. After the agarose gel electrophoresis we have to visualise the amplified product. This is the chemiduct MP where we are going to visualise the amplified product.

Now we have to keep the gel and close the thing. We have to select here application nucleic acids, ethidium bromide, exposure, optimal exposure or we can select manual also. Then we will acquire the images. Now we can find here this is the DNA ladder, this is the PCR amplified product. We can transform it into transform or save this image into JPG.

In this video we have discussed how to set up a PCR reaction and how to analyse the PCR result. We have also shown how to use a thermal cycler and what are the components of thermal cycler. So during all these process we have to take some precautions to get better results. Like all the time you have to keep palmarise engines and your primers on ice. And other thing we have to remember is while running the agarose gel electrophoresis always wear the gloves to prevent any contamination with the ethidium bromide.

So ethidium bromide is a carcinogen. So it is not very likely to cause any cancer but we have to make sure that we are providing this kind of touching ethidium bromide. Now once you set up the PCR you are actually going to face the different types of problems. You are going to face the multiple types of issues like you are not getting the amplifications, you are getting the amplification of the non-specific bands and so on. So how you are actually going to encounter or how you are going to provide the solution to that.

So you are going to do the optimisation. So optimisation would be for the reaction cycles you are going to do a hot star reactions. So hot star means you are actually going to keep the reaction into the machine. For example like and this what we do is we actually keep the tubes into the machine and then we turn on the machine. So what happen is it is actually going to first enter into the stage 1 where it is going to do a 95 degree Celsius for 5 minutes. Then it will enter into the cycles and it is going to do the stage 2 and then it will enter into the stage 3.

On the hot star what you are going to do is you are going to turn on the machine and let the machine reaches into the stage 1. So once the temperature is at 95 degree Celsius then you are going to open the lid and then you are going to insert your reactions. What is the advantage? Advantage is that it will not going to allow the non-specific amplification. So what happen is when you are even if you are keeping the reactions on 37 degree Celsius or in ice some primers actually go and bind some amount of DNA is get you know get access to these primers. So primers actually goes and bind to multiple places non-specifically and that is going to you know initiate the synthesis of the non-specific DNA.

So to avoid that what you can do is you can just go with the hot star. Then sometime you are actually going to optimize this primer sequences you are going to optimize with the different types of enzymes. For example, you may not get the better amplification with the VFU DNA polymerase you may go with the tag DNA polymerase and so on. And then sometime you also going to have the additives. So there are different types of additives what you can use for overcoming the false positives or your actually going to use it for taking care of the non-specific amplifications. For example, if no product is produced or there is no correct size product produced then you can actually be able to use the check the DNA quality.

First thing what you do is you check the template whether the template is correct. Then you reduce the annealing temperatures because you allow suppose you are annealing at 65 suppose you are annealing at 62 degree Celsius. Then you lower down because you are not getting any products. This means you should allow the annealing temperature. So you maybe do it like 59 or you may do it at 58. See whether you are primary annealing sufficiently or not you can increase the time also you can increase the time like from 30 to 50 seconds or so on.

Then you can increase or play with the magnesium concentration because some of the enzyme requires the higher concentration of the magnesium and some enzyme requires the lower concentration of magnesium. So you do not know the what is the optimal magnesium concentration required for tag DNA polymerase to give you the optimal product. Then you can

add the DMSO to add the assay so that it can also give you the some advantage and it going to give you the better amplified product. You can use the different thermostable enzymes as I said in an example like PFU versus if you use the tag DNA polymerase. The tag DNA polymerase will give you the product but PFU may not give you the product because the PFU's Km values are on a higher side compared to the Km values of the tag DNA polymerase which means tag PFU requires the more amount of templates compared to the tag DNA polymerase.

Tag DNA polymerase is more efficient in terms of giving you the product but it is having a problem of the it does not have the proofreading activity whereas the proof PFU has a proofreading activity. So you would suppose to change the enzyme you may actually that may also help in getting the right product. At the end if the thing happens then you should throw out all the primers and you should start the new primers and because sometime the primer also get degraded and 50 get crypted and all that and that is how they may not binding to the correct DNA sequence or they may not binding to the DNA sequence at all. Now if you are getting the non-specific product so you are going to have two condition only no product or you are getting more product then you can actually be able to do many of these things which are you know reverse like for example you can increase the annealing temperature, you can reduce the magnesium concentration, you can reduce the number of cycles and you can also try the different types of enzyme.

We gave you example of PFU versus tag DNA polymerase. Now once you are done with these PCR you are going to analyze the PCR onto agarose horizontal gel electrophoresis right and what you are going to see here is that you are going to get the product. So for example you are going to run the three different reactions you are going to run it along with the marker so that you will be able to know what is the size of the amplified product. So this is actually the negative control so you are not getting any DNA amplified DNA except you are getting the template DNA and then this is the amplified product what you are going to do and these are the primers right what you have used. So remember that in a negative amplification primers are not being consumed and that is why you see a primer whereas in the case of the positive amplification there is no primer right because all the primers are being used up in giving you the amplifications. So once you are done with the analysis of the PCR, the PCR is complete the amplified product is loaded onto the agarose gel and observed under the UV a weak water blank reaction is included to monitor the cross contaminated DNA source as a template.

The percentage of agarose gel depend on the size of the DNA to be visualized generally 0.8 to 1% agarose is used for analyzing the 0.525 kB DNA while a DNA larger size or genomic DNA is visualized in gel with the lower percentage of the agarose. Now why we are doing the PCR we have the you know speed easy of use and sensitivity and robustness and what are the limitations of the PCR as I have discussed right we need a target DNA sequence information you require the fidelity of the DNA applications and so on. So you require the target DNA sequence information so that you can be able to do a primer designing for unexplored ones and you can actually be able to have the boundary region of the DNA to be amplified must be known. Tag DNA polymerase no proofreading mechanism so it is going to have an error of 40% after 20 cycles so short size and limited amount of PCR product is going to be formed then you are going to have up to 5 kB it can be easily been amplified but up to 40 kB can be amplified with some modification

but cannot amplify with the gene which are larger than 100 kB.

So for these kind of large DNA gene fragments you will not be able to amplify with the help of the PCR. The PCR cannot be used in a genome sequencing project because you cannot be amplify the whole genome and that is why you can be able to produce or synthesize that. So these are the some of the limitations so these are some of the theoretical aspect of the PCR what we have discussed in this particular lecture what we have discussed we have discussed how the PCR is been evolved and how the concept of the PCR came by taking the inspiration from the biological phenomena of DNA applications. So people have discovered how the DNA is been synthesized into the cell by the discovery of DNA polymerase by the author Kornberg and his colleagues. But how to modify those events how to modify those things are been always been achieved once the people have discovered the tag DNA polymerase and other kinds of enzymes and then they actually started doing the experiment with the different types of physical parameters and so on so that they can be able to break the double standard DNA into single standard DNA so that they can be able to avoid the role of the helicases, primases and all that. And that is been achieved by mostly by the Carey Mullers and his colleagues and that is how he has developed the technique of polymerase chain reactions.

And in a polymerase chain reactions you require the three events you require the denaturations you require the annealing and you require the elongations. And all these are been performed by the varying the temperature from the 95 degree Celsius to 55 degree Celsius to 72 degree Celsius. And by doing so you can be able to amplify a DNA fragments and then you can be able to get the desirable product. At the end we have also discussed about the different types of the limitations of the PCR and we have also discussed about the troubleshooting part. So with this kind of background information we would like to discuss more about the technical part how you are going to design the primers how you are going to set up the reaction and so on.

So that we are going to discuss into the subsequent lecture. So till then we would like to conclude our lecture here and in our subsequent lecture we are going to discuss some more aspects related to polymerase chain reaction. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 09
Molecular Techniques (Part 2)
Lecture-38 Polymerase Chain reaction (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati and what we were discussing? We were discussing about the different aspects of the molecular biology in the course in this particular course. So, so far what we have discussed? We have discussed about the cellular structures followed by we have discussed about the different types of biomolecules and then we have also discussed about the role of these biomolecules in some of the biological processes such as we have discussed about the cellular metabolism, we have discussed about the apoptosis and death and we have also discussed about the autophagy and we have also discussed about the role of these biomolecules into the cell division and other kind of aspects. Following that we have also discussed about the central dogma of molecular biology and how these biomolecules are very crucial for regulating the different types of molecular processes. We have also discussed about the replications, transcription and translation and we have discussed the replication in prokaryotes and eukaryotes and the other processes also in the both kind of species to give you a comparative study between what happens in the prokaryotic system and what happens in the eukaryotic system. Following that we have also discussed about the some of the molecular techniques.

So, we have discussed about the blotting technique in the previous module and we have discussed about resin blotting, resin blotting and cystin blotting. In the current module we are discussing about the amplification techniques. So, in that amplification technique we are discussing about the polymerase chain reactions and if you recall in the previous lecture we have discussed about the theoretical aspects of these techniques. So, we have discussed how the technique is been evolved by taking the inspiration from the DNA applications.

So, as soon as the author Konberg had discovered the enzymes for the DNA applications and the complete machinery for the DNA applications people have started thinking about modulating the machinery so that it can be done in a better way and that is how we have the scientists have developed the polymerase chain reactions where they are utilizing the different temperatures for performing the different types of events. So, for example, in the initiation you require the melting of the DNA and you also require the addition of the primers and so on and that all been achieved by heating the DNA at 95 degree Celsius for 10 minutes. So, that will be the denaturation steps following that they are actually going to lower down the temperature to annealing temperature so that you can the primers are actually going to go and bind and then they will enter into the elongation phase and there they are actually going to allow the DNA polymerase to go and sit on to the template and they were actually going to do the synthesis of the DNA. They have also discussed that how the different types of components are required and what are the different types of precaution you should take, what is the advantage of the polymerase chain reactions and what is the advantage of using the tag DNA polymerase versus the PFU polymerase and

what are the different types of troubleshooting you can actually be able to do if you encounter the problem of no product or the appearance of the nonspecific products. Now, in today's lecture we are actually going to focus more on the technical aspect of the polymerase chain reaction.

So, in the previous lecture we discussed about the theoretical aspects how you can be able to do all that, but in the current lecture we are going to discuss about the technical aspects so that how you can be able to set up the PCRs, how you can be able to design the primers and so on. So, when we talk about the technical aspect the first thing what comes into the mind is what are the requirement of the polymerase chain reactions and how you are going to set up the polymerase chain reactions. So, as far as the requirement is concerned you require the following reagents, you require the template DNA, you require the primers, you require the magnesium chloride and you require the dNTPs and at the end you also require the tag DNA polymerase. What are I have not written, but you also require a water so that this whole reaction should be of 50 microlitre. So, as far as the template is concerned that template can be vary the amount of template what you are going to add.

So, it could be in the 1 picogram to 1 nanogram for the viral or the short templates whereas, it could be in the 1 nanogram to 1 microgram for the genomic DNA. Remember that the genomic DNA is a large DNA. So, if you are and the PCR is always been depends on the number of molecules, it does not depend on the microgram of the things. So, number of molecules because it is going to amplify based on the number of molecules. For example, if you started with the 10 molecules then it will go after first cycle it is going to be 20 after second cycle it is going to be 40 and so on.

So, it does not depend on the micrograms, it does not multiply the 10 microgram to 20 microgram to 40 microgram to 80 microgram. It multiplies to the number of molecule and that is why it is important that you should take the larger amount because of the larger size of the genomic DNA compared to the viral or the short templates. So, it is you know depends on the moles of that particular template DNA and that is why you see the such a huge difference in terms of the amount of DNA what you are going to use if the DNA is small or the DNA is large. Apart from that you also require the primers. So, you require the two primers one is a forward primer and other one is the reverse primer and the concentration what you require is 0.

1 to 0.5 micro moles of each primer. Then you require the magnesium chloride and magnesium chloride will be 1.5 to 2 millimolar which is optimal for the tagged DNA polymerase. Magnesium chloride is optional it is not required or it is not something which is compulsory because for example, if you use the PFU DNA polymerase then magnesium chloride may or may not be required. Then we also require the dNTPs.

So, difficult concentration of the dNTP is 200 micromolar of each dNTP which means you are going to require the dATP, you require the dCTP, you require the dTTP and then you also require the dGTP right. These are the four nucleotides what is present in the DNA and that is what you require. Then you require the polymerase you can use the tagged DNA polymerase or you can use the PFU. So, tagged DNA polymerase you require into the range of 0.5 to 2 units

per 50 micro reactions.

If you go to other reactions then you are going to require the more amount of the tagged DNA polymerase. And then you also require the water and as well as the polymerase buffer also. Now, one of the most crucial component in this whole reaction is the primers right. The better you remember that when we were discussing about the troubleshooting we gave you the many types of excuses many types of options how you can be able to improve the product and as well as the accuracy of amplification. And one of the major reason why you are getting the non specific amplification or why you are getting the no amplification is because you have designed the primers which are not correct or you have designed the primer which are not serving the purpose of working as the template or working as a attachment point.

Because remember what is the job of the primer? Primer for example, if this is the 5 prime to 3 prime right and if you add a primer right because you remember that the DNA polymerase has a limitation that it cannot start the DNA synthesis from the nascent DNA. It cannot start from the just by looking at the nucleotide. It requires the some kind of attachment point and that is why you require a primer. So, if you add a primer which is so, then it is going to bind as per the you know the what centric base pairing information right. For example, if you have A here then it is going to be T here and so on.

So, and then this last nucleotide is actually going to provide a starting point for the DNA polymerase and that is how the DNA polymerase will come sit here and it will go into start synthesis of the DNA right. So, if the primers are not good they are sitting in a very random fashion like if they are sitting somewhere here or if they are sitting somewhere here or if they are not sitting at all on to the target DNA they are sitting on to somewhere else. So, for example, if you have another DNA and it may be sitting somewhere here then it may be initiate the synthesis of that DNA because tag DNA polymerase does not know which DNA it supposed to synthesize. So, in that case it only goes to know that wherever I will find the attachment point wherever I will find the binding of the primer I will extend that particular DNA. So, that is why the accuracy and the primer of the DNA or the primer what you are going to design is the most crucial component of the polymerase chain reactions and that is why we are in we are going to discuss in detail about how you can be able to design the primers and how you can be overcome these kind of limitations.

So, as far as the primer is concerned primer is a short DNA stretch that serve as a starting point for the DNA synthesis. In PCR the two primers are required to bind to the each of the single standard DNA you will obtain that after by denaturation flanking the target sequence. These are called forward primer and the reverse primer. They are the pre-mermer R sequence complementary to the sequence in the template DNA where they are supposed to start the synthesis. So, that the in the PCR the primers are the made up of the DNA.

So, they are actually deoxyribonucleic acid whereas, in the case of the DNA replication they are you have seen that we are actually going to use the primers which is being seen synthesized by the primase it is actually RNA. So, that is very very different because in that case you actually

have the machinery which actually is going to remove the RNA and it is actually going to replaced by a short stretch of DNA by the some kind of enzymatic system right. But here you do not have that and that is why you are actually going to put the DNA single standard DNA as the primer right. And you require the two primers forward primer and the reverse primer. Forward primer is for the synthesis of the leading strand or reverse strand for the reverse primer for the synthesis of the lagging strand.

Although I just said in the previous lecture also and I am going to clarify today also that there is no leading and lagging strand because there is no delay in the synthesis of the one strand versus another strand. It is actually going to start simultaneously right. Suppose you have the DNA right once you are actually going to have the denaturation it is actually going to have the two different independent templates and that is how you are going to have the independent finding of the forward primers, finding of the reverse primer and that is how it is actually going to do the synthesis of both the strands simultaneously. Just for the sake of making it analogy with the DNA replication I am using the term as leading and the lagging strand, but there is no leading strand there is no lagging strand in the case of PCR. Now as I said in the beginning also that the primer designing and the quality of the primer is very very important and that is why there is a set rule and set guidelines how you can be able to use the primers or how you can be able to use those guidelines for designing the primers.

Primers so, first is what is the length of the primer? So, oligonucleotide between 18 to 25 bases is a ideal length which is long enough for the adequate specificity and short enough for primer to bind easily to template at the annealing temperature. This means 18 to 22, 24 is the ideal length if you go little shorter to this then the DNA will not going to bind to the specific site. It is actually going to lose its specificity because that small stretch of DNA probably will find a compliment DNA into the multiple places. For example, it may actually be you know 4 nucleotides suppose you make a primer of 10 nucleotides right if you make a primer of 10 nucleotide then that 10 nucleotide by chance could be present in any part of the genome and that is how it is actually going to compromise the specificity. So, that is why it has to be a certain length so, that that particular complete stretch may be unique for that particular stretch of DNA, but it may not vary to other site.

Now the other question is if you make the very long DNA right if you make a very long primer then the primer will actually going to have the many kind of problems right it may have the secondary structures in within the primer then and it may also have the problem of provide giving you the non specific amplifications because very big like long like for example, if I make a primer of 50 base pair then I am actually you know giving more sequence and I am giving the that particular type of sequence which are not required. So, that those additional sequence additional base pairs are actually going to form the different types of secondary structures and the secondary structure could be problematic in terms of allowing the primer to anneal. The second point is that this long stretch of DNA may be actually finding difficulty in annealing very you know precisely to that template for example, it may anneal like this for example, it may anneal like this right. So, some of the DNA is finding the complementity the other strand is finding the difficult to anneal by the time it actually finding the you know the complementity this portion has already

been annealed right. So, our other example is that it may actually anneal here small stretch and the rest is leaving like this and then goes although it has the complementarity here, but the kinetics of binding of this region is slightly faster than kinetics of this that is why it may not get chance to bind and then this portion also has a faster kinetics.

So, it actually goes and bind. So, that is why you are going to have this kind of you know loop or hair pins and that may or also going to affect. The other point is also the cost right because this is actually going to increase the cost because if you are designing a primer of 50 base pair. So, it because the primers are actually going to be synthesized under the in vitro DNA synthesis reactions and that is going to be supplied by a commercial vendor. So, if you are unnecessarily putting another 26 base pair then you are supposed to pay for those synthesis and you are supposed to pay for those primers.

The second point is about primer melting temperature. So, primer with the melting temperature in the range of 52 to 58 degree Celsius generally give the best result. The GC content of the sequence gives a fair indication of a primer Tm. The two primers should be prepared in such a way that the Tm difference should not be more than 2 degree otherwise it will give you the poor annealing efficiency. So, primer first point is going to take care of the annealing.

So, first point is going to take care about the complementarity and it is going to be good for but second point is very very important. So, second point is that if you are going to have the annealing temperature in the range of 52 you know 52 to 58 then when you are going to lower down the temperature the primer will have the better affinity and it is going to anneal right. And that you can actually be able to achieve by keeping the very significant amount of the GC content. I will discuss in detail why you supposed to have the GC content because the GC content is also very important in providing the stability to that interaction because you know that the G is interacting with the C with the 3 hydrogen bonding and it is stronger compared to the AT interactions right. So, AT is having 2 hydrogen bonding and then you can actually be able to calculate the melting temperature.

So, there are 2 formulas if you are having a primer length which is less than 14 then you can use the Tm formula of this like 4 degree into number of G and C in the Tm in the primers plus 2 into number of A and T in the primer. But if it is length more than 13 nucleotide then you can use the this formula for calculating the Tm values. Now the third point is about primer annealing temperature. So, you require the 2 parameters one is primer and melting temperature and another one is the primer annealing temperature. So, too high Ta will produce the insufficient primer tabular hematization results in the lower PCR product yield whereas, the too low will lead to the non-specific product caused by a higher phase pair.

And how you are going to calculate the primer annealing temperature? So, you can actually be able to calculate using this particular formula and Tm for where the Tm is the primer is the melting temperature of the primers and Tm of the product is the melting temperature of the product right. So, that means, that the synthesized product right. So, the synthesized DNA because the synthesized DNA is also going to have the affinity for the template right. So, that also you can

be able to calculate the melting temperature for them. So, you can calculate the melting temperature for the primers you can calculate the melting temperature for the product and then you put it into this particular equation and it is actually going to give you the annealing temperature for the primers.

Then the fourth point is very very important then you supposed to have the GC content the number of G's and C's in the primer as a should be in the range of 40 to 60 base pair. And the GC content is very important because it is actually as I said you know in the previous slide itself that G is making a base pair with C with the 3 hydrogen bonding compared to the A to T because T A to T is always responsible for having the 2 hydrogen bonding. So, because of it is strong the binding of the primer to the template is going to be strong right it is going to be very strong and that is how it will allow the DNA polymerase to go and sit and do a efficient synthesis. It is going to give you the efficient initiation and once the initiation start then there will be no doubt that you are going to get a PCR product. But if this is very weak for example, if you have lot of AT and G those kind of sequences then what will happen is that the primer will be binding primer will still be binding because the 2 base pairs are still good enough right.

But they are not good enough right they are not efficient they are not going to provide the efficient initiation and because of that the primer RNA polymerase or DNA polymerase may start, but it may take time because it has to sit on this double standard DNA and then it has to actually extend this strand right. So, this for this reason only it has to be a GC content. Now apart from that you also should have the GC clamp right. So, it is not important that you should have a GC content you also should have that the corner and the corner you should have the G G and C kind of sequences. So that the corner should be intact ok even if this region is not having the lower affinity it does not matter because this portion can be taken care.

But if the lower if this stretch like the corner of the DNA is also coming off right it is because of the lower affinity or some other reason then it is will be a problem. So that is why you should have a GC clamp on the corner and as the GC forms a strong bond with then the AT the number of GC content at the 3 prime end of the primer should not be more than 3 otherwise it will result in a non specific binding at the region where the G and C are abundant. So, this is this is very important that that you should have a GC clamps to hold the primers. And then we have you are supposed to design the primer. So, I have taken an example of this you know.

So, for example, if this is your gene for which you are going to have the amplification right or you want to design an amplification. So, apart from the GC content you are supposed to have the you are not supposed to have a GC clamp for example. So, if you have a primer where you have the you know the G G G like that sequences at the corner then it is actually going to have the problem because it may actually flip on to this and it actually can make a loop like this. So, it is going to form a clamp. So, as the GC forms a stronger bond than the AT the number of GC content at the 3 prime end of the primer should not be more than 3 otherwise it will result in a non specific tight binding at the region where the GC are abundant.

So, this is very important that you should not have the GC clamp you should have not have the

more than 3 nucleotides at the corner because otherwise the binding is going to be very strong and it will result into the non specific tight binding at the region where the GC are abundant which means it is actually going to guide the binding of the primer into any sequence where you have this complementary sequence. And in so because this sequence does not match this sequence is not matching with this, but this sequence is matching and because it is at the corner it will allow the DNA primer to go and sit. So, this is a non specific template this is not a specific template, but it will still be get amplified and that is why you are going to have the non specific amplifications. Now, let us take an example of this double standard DNA. So, this is the double standard DNA which you are interested to do a PCR.

Now what we are supposed to design a forward primer and you are supposed to design a reverse primer. When you want to design the primer what you require is you require an information about the restriction enzyme because most of these primers are going to use for cloning purposes. So, when you design the primer you are also going to ask a question whether I want the primer for sequencing purpose. Remember when we were discussing about the Maxine Gilbert method we are saying that you can actually be able to use the primer that is how you can be actually you know design the four different types of reactions A reaction, G reaction, C reaction and D reactions and so on. So, whether you are designing a primer for the sequencing purpose or whether you are designing a primer for the cloning purpose.

90 percent people are designing a primer for the cloning purpose. So, if you are requiring if you are designing a primer for sequencing purpose then the requirements are different then you do not require the restriction enzyme, but for the cloning you require following things. You require a restriction enzyme because this is the restriction enzyme what you are going to use for cutting the fragment and then you are getting going to paste it into your vector of interest. So, you require the information of the restriction enzymes then you also require the you know the sequence information sequence of the template right and as I said you know that is the one of the limitation of the PCR. So, for example, this is the stretch of DNA and I want to clone this into PET23A ok.

This is just an arbitrary example it could be any other vector. So, it could be a one of the vector which is where I would be interested to clone this right. Now, the first thing what I have to see is I have to go through with the PET23A multiple cloning sites and I have to look for the restriction enzyme what I suppose to use right. Restriction enzyme what I can use right. So, I will have a list of restriction enzyme which are present in the NCS and then I have to also look at the sequence of my template right and I have to list out the non cutters right those restriction enzyme which are not cutting the sequence right and then I can use non cutters from this list and then I can choose the enzyme.

So, for example, in this case I have chosen that I will use the BAMH1 for the fiber primer and I will use the XHO because looking at the MCS you will be able to know which one which enzyme is in the front and which enzyme is on the 3 prime end. So, I can use the BAMH1 in the 5 prime end and I can use the XHO in the 3 prime end ok. So, this means the gene is going to be placed between the BAMH1 and the XHO1 ok. So, this is your DNA this is the DNA you are going to have the restriction enzyme in the front you are going to have restriction enzyme in the bottom.

So, that when you are going to digest this with BAMH1 versus XHO you are going to have the BAMH1 site here XHO site here and then you are going to have the similar kind of cutting for the vector also and that is how it is actually go and insert into the vector.

Now for this what you require is you are going to have the simple thing ok. What you can just do for forward primer it is not an issue forward primer what you have to do is you have to just take the 3 prime end whatever the sequence you require. So, you can take the 15 to 20 nucleotides from the 3 prime end right you can just take the in fact what you can do is you can just take the 5 prime end ok and then you can just add the. So, you can take the 5 from the 5 prime end ok and then you put the restriction enzyme in the front ok. Now this is your primer ready this is your forward primer ready ok and you can write the name you can just give some name like F1 or you can say F1 X gene like that ok because it is important that you should give the better name.

So, that you should be able to you know get this primer at a later point. Then for the reverse primer what you are going to do is you are going to generate the 5 prime end of this. So, you are going to write the 5 prime end of this and you are going to write the 3 prime end to this and you are going to write the complementary sequence to this and then you are going to do the same you are going to put the restriction enzyme here you are going to put the restriction enzyme 1 that is the Bamach 1 and the other one you are going to use restriction enzyme 2. And you can finally, you what you are going to do is you are going to do lot of quality testing you are going to look at whether the primers are annealing with each other or not whether there is a primer dimer is forming or not and so on. So, all these I have done or we have done it in a demo video.

So, that the students are actually going to explain you how you can be able to do the design primer manual method and as well as the software based method because there are so many software available where you are just going to put your gene and they will actually going to give you the you know the simple examples of the these are the potential forward primer and potential reverse primers. So, you can get the readily you can get the sequences. So, for designing primers first you have to identify the region of interest your region of interest which you want to amplify from any vector or any sequence. So, in second step you have to identify non cutters. There are various software available but we can use New England Biolabs NEB cutter version 2.

0. After identifying non cutters you have to select a suitable vector in which you want to integrate this amplified region and suitable restriction sites. You will get suitable restriction sites from non cutters. After that you can go for designing forward primer. So for understanding purpose I gave this sequence.

So I am using this sequence. I will use this sequence to design the primers and analyze the primers. So this is the whole sequence but I don't want to amplify whole region. I want to amplify the letters the sequence which is highlighted in green. So I want to amplify starting from here to here. So now the question arises what are the non cutters? So you want to amplify this region and integrate into another vector for that you have to identify which are non cutting restriction enzymes.

So what I will do I will copy this sequence into NEB cutter and identify what are the non cutters. So I just copy the sequence paste here and I will ask submit. So it will analyze the sequence and give non cutter. These are the enzymes cutting inside the sequence.

But we are interested in which are non cutters. So that means you can see here non cutters. So just click here. It will give a number of enzymes which will not cut inside the sequence. So once getting this list we have to identify in which vector you want to integrate your amplified region. So for that purpose so I have selected for easy of understanding I have selected PED 23A vector.

So you can see this is the vector map. So this is the 5 prime side. This is the 3 prime side. N terminal and this is the C terminal side. N terminal means forward primer. C terminal means reverse primer.

So I can use BamH1 in forward primer and Xho1 in reverse primer. This is the detailed map. So I have identified two restriction enzymes that is BamH1 and Xho1. So I can use these enzymes in forward primer and reverse primer. So after identifying restriction enzymes and the vector we will go for designing forward primer.

So I will take this sequence. I want to amplify from here to here. So I will copy this sequence here. So for designing forward primer it is very easy. You have to take the sequence whatever you are getting up to 15 to 20 bases you can take as this. So if you want to insert a restriction enzyme.

Suppose I want to insert a restriction enzyme. This is the sequence as it is given from this whole sequence. So I want to insert a restriction enzyme that is BamH1. So this is the sequence for BamH1.

Here it cuts. So I can use this sequence here. So this is the this is our restriction enzyme here it will cut. So we cannot simply Q like this. So there should be some more bases extra bases.

We have to add in the 5 prime side. So I will use. So this sequence I will use. So now this is 5 prime to 3 prime side. This is our forward primer is ready. So after designing this forward primer we have to analyze this sequence.

So this primer. So what I will do is I just copy this sequence and I will use aligo analyzer software which is specially designed for this purpose only. I will paste the sequence. Just ask analyze. So here also you can see there are so many options are there like you can analyze hairpin loop, self-dimer, hetero-dimer.

So these are the general details. What is the length and GC content, melting temperature, molecular weight. So these are normal details. I will go for hairpin loop. Is there any hairpin loops? So we can see there are a number of hairpin loops.

We can see different structures predicted by the software. So if you want to explore this thing, you can explore only two bases, two bases it is forming and the delta G value is minus 0.43 kilocalories per mole. So this is fine. Up to minus 10 kilocalories per mole is fine. Suppose hairpin loops are broken during the amplification process, but above that, above minus 10 kilocalories per mole cannot be broken.

So in that case, what we will do? Either we redesign the primers or we will add 5%, 1% B10 or 5% DMSO. These are these chemicals disrupt these loops so that the amplification will be fine. So next I will analyze for self-dimer. Is there any self-dimer and what is the maximum delta G? So this is forming continuously 5 bases.

It is because of the restriction sites. So those are restriction sites. Those homodimers forming due to restriction site can be broken. There is no issue. But other than that, this is also because of restriction site. But other than that, we have to look carefully. So is there any continuously 4 or 5 bases forming this homodimer, then it is very difficult.

These interactions can be broken easily. So here are some of the consecutive base pairs. These are very weak interactions. So they can be broken. So other than that, there are no significant self-dimers. So this sequence can be used. And for heterodimer, predicting heterodimer, you need complementary sequence with reverse primer, like reverse primer you need.

So that we will discuss later on. So we got our forward primer here. So it is very easy to generate forward primer. But in case of reverse primer, it is somewhat difficult because not in terms of predicting things, it is somewhat tricky. So what I am saying is, here we have sequence. So in case of forward primer, we just take an as-t sequence, 15 to 20 bases as-t from sequence. But here we have to take complementary sequence, not 3 prime to 5 prime or 5 prime to 3 prime sequence.

We have to take complementary to this one. Say this is the sequence we got from here. So what is the complementary to this one? So just I will add here. So this is the complementary to this particular sequence. So as you can see, this is, we have to keep from this direction, 5 prime to 3.

So I will take like this. So what we have to do is, we want to insert a restriction site here. So we can insert a restriction site here directly. So in reverse primer, we wanted to insert Xho1 site. So this is the restriction site.

As usual, we can use, we have to insert T here. So this is the restriction site we added. We can add flanking regions in between flanking bases before this restriction site. So now we got our reverse primer. So we have to go through the same procedure like what I have shown in case of forward primer. So just I will copy and paste here and analyze the reverse primer.

So is there any hairpin loop? Only one hairpin loop that is within the range of delta G. So there is no issue and self-dimer. So we can see here continuously 4 bases are forming. In this case, we have to either change the sequence or remove some of the bases. We can ignore those restriction, those dimers forming through the restriction site.

So next heterodimer, we have to analyze. For heterodimer, we need forward primer. Just copy, paste here and calculate. It will give is there any heterodimer.

This is because of restriction site. This is also because of restriction site. This can be broken. Those which are at the end of the sequence, they can be broken but which is in the middle if those bases are middle, it is very hard to disrupt those interactions and our amplification will be not good. So there is no amplification literally. Other kinds of interactions will be broken easily.

These are weak interactions. So this is how we can prepare, design the primers and analyze the primers. We have done all these processes for designing forward and reverse primers. But instead of doing manually, we can do it online. We just have to submit the sequence and it will return the forward and reverse primers. These are some of the tools available online for freely but there are commercial tools also available like algo 7 vector NTI primer firmware.

So if you are interested in these software, you can just go through these sites and submit your sequence. You will get your primers. In this video, we showed you how to design forward and reverse primers and how to analyze for non-cutters and what are the restriction sites we can use based on non-cutters and how to integrate our occupied product in which regions like we have to observe the vector if we want to integrate our gene of interest. So hope this will help you to advance your work. Now once you design the primer right, so once after this demo probably you will be able to design the primer on your own and you can be able to test these primers under the in silico PCR. So, you can actually be able to do that by using the some of the in silico tools what are available on to the web.

So I hope you could have been understood the process. Now let us think about the what could go wrong and what are the different problems what you are going to face. So one of the primer one of the major problem what you are going to face is the primer secondary structures. So primer secondary structure arise as a result of intra or intramolecular attraction between the primer or the other primer which eventually reduce the yield or amplification at the availability of single standard primer will be limited to the PCR. The various types of primer secondary structure are as follows for example, you can have hair pins, you can have dimers, you can have repeats and the runs.

So hair pins are the loop structure formed by the intramolecular interaction within the primers. Optimally a three prime end primer with a hair pin with the delta G of minus 2 kilocalorie and an internal hair pin with the delta G of minus 3 kilocalorie per mole is tolerated generally. Then we have a primer a primer is a structure formed from the double standard like structure which is formed by the intermolecular interaction between the two primer. If the interaction is formed between the two homologous or same sense primer it is called as self-dimer whereas, if the

interaction is formed between the two different primer then it is called as cross primers. Optimally a three prime end self-dimer with a delta G of minus 5 kilocalorie per mole and an internal self-dimer with a delta G of minus 6 kilocalorie is tolerated generally. So, dimer is going to be formed between the two primer for example, this is a forward primer and this is going to be a reverse primer.

So, they may actually have a complement between them and they may actually have like binding like this. So, they may have a binding like this they can have multiple options like if they can have a binding like this and so on. So, this is the small stretch if it is having a small stretch and if it is having a delta G in the range of minus 5 kilocalorie to minus 6 kilocalories then it can be broken down when you are going to have a very high temperature, but if it is a very tight binding and if it is having the. So, this is the primer dimers between the primers it could be also that forward primer itself is binding to the forward primer. For example, you can have a forward primer binding to another forward primer because you have some sequence which is complementary to each other or you can have like this I can have complete binding right in the middle actually and so on.

So, this is can be well tolerated also if you have you know the bind the delta G in the range of 5 to 6. Then you also have the repeats and runs. So, repeats are the consecutive occurrence of dinucleotide whereas, runs are continuous stitches of signal nucleotide. A maximum number of repeat and run accepted is the 4 dinucleotide and the 4 base pair respectively. So, repeat and runs are also going to have the similar kind of problem that they are actually going to create the hairpins and all those kind of problems.

Then you can have the primer template homology. The primer should be designed in such a way that there should be no homologous within the template other than the target site. This will result in a non specific binding and the amplification. So, this is just a few examples of the primer secondary structures. So, you can have the for example, if this I have synthesized a forward primer and I have synthesized a reverse primer and this is the primer sequences.

So, then I have to analyze them whether they are forming the loops or hairpins and those kind of errors. So, for example, in this case what you see here is this is the actually the hairpin what is being formed and it is very strong because it has been bound by the two different interactions. So, this is the GC interactions and this is the AT interaction and that is very very problematic because and in if you want to have this and why it is happening because it has a very strong you know the intermolecular primer-dimer formations. So, if I have this and you will see the delta G delta G is in the range of 9.

47 which is above the 0.6 the number above to the 6. This means this is cannot be broken down this is very very strong it cannot be broken down even if you have increase the temperature. So, in that case I have an option of either using the some other stretch of the of the template DNA or I have to modify the primers. For example, I have modified the primer and then now I have broken down that and what you see here is delta G is still you know in the range of minus 2 right. So, it is still there is a interaction it is still there is a 3 nucleotide what is binding to the corresponding you know the templates, but this is well tolerated this is going to be interaction this is going to be

bind and this kind of interaction you do not have to worry about. So, this is the way you supposed to know vary the you are supposed to analyze the primer sequences you are supposed to check for these kind of things and there are software available there are software available for doing this and that is how you can be able to do the primer designing and other things.

So, this is all about the technical aspects what we have discussed we have discussed about the primer designing what kind of troubleshooting you are supposed to do and what are the how it is actually going to impact the PCR amplification and other things. So, with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you. Thank you.

Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module - 09
Molecular Techniques (Part 2)
Lecture-39 Real Time PCR

Hello everyone. This is Dr. Vishal Trivedi from department of bioscience and bioengineering IIT Guwahati. And in the course molecular biology we are discussing about the different aspects. So in this current module we were discussing about the polymerase chain reactions. And if you recall in the previous two lectures we have discussed about the polymerase chain reactions.

We have discussed about the how the PCR can be evolved from the very crude you know thermal cycler to the very very refined and sophisticated machines and how you can be able to design the or perform the PCR. So we have discussed about the primary designing we discussed about the how you can be able to isolate DNA RNA and all that and how you can be able to perform the PCR. And then at the end when you are done with the PCR you can be able to analyze the PCR also. Now in today's lecture we are going to discuss about the real time PCR.

And the real time PCR is a slight variant of the traditional PCR what we have discussed so far. And it has additional advantage as well. So in the current in this particular lecture we will discuss about the real time PCR and how you can be able to perform the PCR into the your laboratory. So for the first question comes what is the real time PCR and what is the you know advantage of real time PCR. So the first question comes what is the real time PCR and what is the advantage of real time PCR compared to the normal PCR.

So the real time PCR as the name suggests is actually a PCR where you are going to do the amplification cycles and at the same time you are also going to monitor the product. So real time PCR actually combines the amplification of the target DNA and its detection in the single step that is it collects the data throughout the amplification in a real time. Remember that when we were talking about the polymerase chain reactions we were having the four different steps. So we were having the denaturations, annealing, then extension and then that cycle continues. So after at the end of every cycle you are supposed to you know in the traditional PCR you will not be able to you know know what is the amount of DNA is being produced whereas in the real time PCR you are actually going to see what will be the amount of product is being developed.

So real time PCR combines the amplification of the gene of interest that is the you know traditional PCR what we are doing also with the intensity of the fluorescence. So this intensity of the fluorescence is directly proportional to the product what is being formed. Remember that when we were talking about additional PCR we are saying that it is you know we are going to run the 25 cycle, 35 cycles and so on. So after end of every cycle it is going to tell you what will be the amount of fluorescence and you can easily correlate that fluorescence to the amount of

product what is being formed. The movement at which the target amplification is observed first mainly define the reactions the time period at which the fluorescence intensity exceed the background fluorescence intensity is called as the cycle threshold or CT values.

As a result a large amount of target DNA the fluorescence signal appear more quickly resulting into the lower CT values. The people sometimes people are not saying the real time PCR they are also using the different words such as the kinetic chain PCR or the quantitative PCR. Quantitative PCR is also been called as qPCR. So there is a slight difference between the traditional PCR versus the real time PCR. In a traditional PCR you are not been able to you are only going to know the end product.

That means after every 25 cycles what will be the amount of DNA is been produced. Whereas in this case you are actually going to see the amplification after every cycle and that is why you it is actually going to give you the real life you know real time monitoring of the product. Now the question comes what will be the advantage of using the real time PCR. So why there is a real time PCR preferred over the fundamental PCR. So the first option is it is going to be a quantitative analysis which means the real time PCR is going to allow for the quantitative measurement of the DNA or the RNA what is present in the sample.

And it is actually going to provide the quantitative results which means it is actually going to tell you okay 10 microgram of DNA is been produced after first cycle 10 cycle fourth cycle like that. Number two is the speed okay. So real time PCR provide the result in a real time hence provide result in a short duration of time. Fundamental PCR requires the post reaction analysis which can be time consuming and so in a fundamental PCR you are actually going to you know run the product onto agarose gel and then you are actually going to you know do the you know transimulator and then you are going to observe the DNA. Whereas in this case it is actually going to give you the real time PCR you know information about what is the product been formed.

Number three the real time PCR it is more sensitive. So real time PCR is more sensitive and can detect the low copy number of target DNA due to the continuous monitoring of the sample. The fundamental PCR is less sensitive because it is actually going to give you the end product. Number four real time PCR is reduced contamination risk. So the closed tube system of the real time PCR reduces the contamination risk due to the minimum post PCR handling.

It is the main concern with the fundamental PCR. Then data accuracy since this is quantitative the data is going to be more accurate. So real time PCR provides the precise and the accurate data throughout the quantification of the frozen signal. Real time PCR is more susceptible to the variation because it depends upon the you know when you are going to run the DNA onto the agarose gel depending upon the many factors your quantitative quantitation could be wrong. Number six it is high throughput.

So real time PCR can be easily automated for high throughput applications making it suitable for the large scale testing. Remember that real time PCR people were using very you know often in the case of COVID testing right because of this feature only that you can actually be able to

you know do like 50 samples 100 samples 2000 samples and so on because it is it can be automated with the help of the high throughput things. And number seven you can actually do a multiplexing. So real time PCR enables the simultaneous detection of the multiple target in a single reaction which is very very difficult to do with the traditional PCR. So because of these seven advantages people are doing the real time PCR because it is giving you more information about your biological samples than the traditional PCR.

Now what is the principle of the real time PCR? So the principle of real time PCR is to quantitatively measure the amount of specific DNA or RNA target in a sample by continuously monitoring the amplification process in a real time fluorescence. This technique leverage the natural ability of DNA to produce a fluorescence signal when it is amplified. There are two common methods which are being used. One is cyber green dye method and other one is called as the TaqMan probe method. So the cyber green method in this method a fluorescent dye such as cyber green is added to the PCR reaction mixture.

Cyber green binds to any double standard DNA and this generates during the process as the DNA target amplifies more double standard DNA is produced leading to an increase in fluorescence. This rise in fluorescence is directly proportional to the amount of target DNA. However it is important to design the specific primer to ensure that the fluorescence signal correspond to the specific target. So this is exactly what is going to happen. When you are doing a PCR reactions you are going to do the denaturation, you are going to do the annealing of the primers, then you are going to do the extensions and after every cycle you are actually going to have the more amount of double standard DNA.

Remember that after if you start with the one amount one DNA molecule then after first cycle you are going to have so after first cycle you are going to have two copy of DNA and again you are going to have four copy of DNA and so on. And as the amount of DNA is going to increase it is actually going to bind the cyber green. So cyber green is a dye which actually goes and binds to the double standard DNA. So as soon as you have the one amount of DNA its cyber green is going to give you the X fluorescence. If the two DNA molecules are being formed then the two molecules of cyber green is going to bind and that is why you are going to get the 2 X fluorescence and you are going to get 4 X fluorescence and so on.

And that is why you are actually going to get the fluorescence signal proportional to the amount of DNA. And that is what it is going to happen. So once you are going to do the DNA synthesis after DNA synthesis you are going to have the two amount two strand of DNA and the cyber green is actually going to intercalate into the DNA stuck DNA and it is going to give you the fluorescence. So this is the first method. The second method is the TaqMan method TaqMan probe method.

So TaqMan probe method is more advanced method. So in this method apply a specific sequence specific fluorescent probe like TaqMan probe that is designed to bind the target DNA. When the probe binds to the target during the DNA synthesis the DNA polymerase cleaves the probe releasing its fluorescence reporter. The increase in fluorescence is directly proportional to

the amplification of the target sequence. This method offer high specificity because the probe only fluorescence when it is bind to the intended targets.

The real time PCR instrument set a predefined fluorescence threshold and the cycle at which the fluorescence signal surpasses the threshold that is recorded as the cycle threshold or the CT values and the CT value is very very important. So in a TaqMan method what you are going to do is you are going to do annealing. So when you are going to do annealing the your TaqMan probe is actually going to go and bind and when there will be polymerization and the strand displacement what will happen is that this fluorescent dye is actually going to be cleaved by the DNA polymerase and it is actually going to be released into the signal released into the solution and that is actually going to be detected by the detection system and is actually going to be proportional to the amount of DNA what is been synthesized. And this is exactly what you are going to go. You are going to see the increase in the fluorescence signal and then it reaches to the threshold and then it actually goes and reaches to the saturation.

So when it reaches to the saturation then it is actually going to called as the cycle threshold or the CT value. So this is actually going to be CT values. And so this is going to be the maximum fluorescence what it is actually going to achieve. And the CT value is inversely proportional to the amount of target DNA or RNA in the sample like the lower value CT values indicate a higher concentration of the target sample. This means when it reaches to this threshold values it is actually going to reach this point if you are going to have the.

So it takes 10 cycles it can take 15 cycles it can give 30 cycles. So more number of cycles it reaches and crosses the threshold cycle threshold values that means that the lower the concentration of the DNA or RNA was present in your sample. So it is inversely proportional to the amount of DNA then the amount of time it is going to take to reach to the threshold values. The specialized software which process the you know which process the flows and data to calculate the CT value and a standard curve is generated using the unknown concentration of the reference DNA and RNA. This is standard curve enables the determination of the initial amount of the target sequence in the sample.

Now if you want to perform the real time PCR you require a set of reagents you require the set of machines and so on. So let us see what are the things you required. You require a thermal cycler and this thermal cycler is different than the thermal cycler what we were using for the traditional PCR because this thermal cycler should have a detection system so that it can be able to detect the fluorescence. So you require a thermal cycler and you also require a fluorescent measurement system so that it and you can integrate that into the machine so that while it is amplifying the reactions it can also detect either the cyber green method or the TaqMan probe method. Then you require the DNA or RNA as a template right mostly people use the RNA as a template because that RNA is the most desirable thing what you can do but you can also do the DNA also and the target or DNA which can you want to quantify or amplify then you require steep primers you require the forward and the reverse primers you requires the enzymes so you require the two different enzymes you require the DNA polymerase and you also require the reverse transcriptase.

So DNA polymerase is responsible for the synthesis of new DNA strand during the PCR so it is exactly the same as what we were using for the traditional PCR and then you also require the reverse transcriptase so you can use to quantify that RNA which is done by converting the RNA into cDNA and before the process of amplifications. Then you require the fluorescent dye so either a fluorescent dye that binds to the double standard DNA like cyber green or you can use a specific fluorescent probe that bind to the target DNA generating a signal during amplification that means you can use the TaqMan probe method. Then you require the detection system so you require the real time detection system so this system requires the fluorescent measurements during PCR cycle commonly cyber green or the TaqMan detection systems are used. So in many of the time what happen is that this real time detection system and the thermal cycles are actually going to be integrated into each other and that is why it is going to be called as real time PCR machine. And then you also require the PCR tubes or the plates.

Now before moving into how you are going to perform the real time PCR you also should understand how the reverse transcriptase is going to work. So reverse transcriptase are like RNA dependent DNA polymerase so it is exactly the reverse what the RNA polymerase is doing. So RNA polymerase is doing the RNA synthesis from the DNA reverse transcriptase is doing the DNA synthesis from the RNA falls into the category of polymerase enzyme which helps in the generation of cDNA taking RNA as a sample. Reverse transcriptase can perform three activities it can perform the RNA dependent DNA polymerase that uses the single standard DNA as a template to generate the cDNA it can do the RNase H endonuclease activity so it can degrade the RNA strands of DNA-RNA hybrids and it can do the DNA dependent DNA polymerase. So converting the single standard RNA probe into a double standard DNA all these molecular techniques such as real time PCR, RTQ PCR, cloning of cDNA, RNA sequencing frequently uses the reverse transcriptase at requires conversion of RNA into DNA.

So reverse transcriptase has two activities polymerase active site and endonuclease active site and it is actually going to have the three different types of enzymatic activity RNA dependent DNA polymerase, RNase H endonuclease and the DNA dependent DNA polymerase. And exactly this is what it is going to do it is going to take up your single standard RNA molecules then it is actually going to do the polymerization reactions so when the reverse transcriptase is going to work it is going to use the you know the it is going to synthesize the DNA and it is going to produce a RNA-DNA hybrid and you know that reverse transcriptase is also going to have the RNase H activity so it is actually going to degrade the RNA part so it is going to degrade the RNA from the this RNA hybrid and leaving the single standard DNA template and then this single standard DNA template is also going to be converted into double standard activity by the DNA dependent DNA polymerase activity and at the end what you are going to do is you are going to get a double standard DNA template from the single standard RNA template and this double standard RNA template is going to be called as the cDNA or the complementary DNA. There are many more methods through which you can be able to prepare the cDNA but most of the people use the reverse transcriptase because it is easy and it is straightforward. Now once you prepare the cDNA right you can be able to perform the real-time PCR. So there are steps which are involved into the real-time PCR like you are going to have the first step that is the sample

preparation.

Sample preparation means you are going to denature the cells suppose you started with the cell or you have started with the fluid or whatever so you are going to crush the cells you are going to denature the and you are going to extract the RNA or DNA whatever the target molecule you want to quantify and then you are going to extract the DNA or the RNA. If you are going to extract the DNA then it is directly going to get into this reactions. If you are going to start with the RNA then it is actually going to have the additional step of converting the DNA into RNA into DNA with the help of the reverse transcriptase. Then you are going to have the primer and the probe designing so it's going to you are going to design the primers and probe and all that and then you are going to set up the reactions. We are going to set up the body real-time PCR reactions where you are going to have the template DNA or cDNA.

You're going to have the cyber green you are going to have all other kinds of you know the reagents what you are going to add then you are going to set up the amplification cycles you are going to do the flows and monitoring and then you are going to have the threshold detections and then ultimately you are going to do the data analysis. Now each step in each step like for example, in the step number 1 you are going to have the sample preparation. So, you are going to extract and purify the RNA or the DNA whichever the molecule you want to you know isolate or detect from the biological sample both quantify and quality are important factor for the accurate results. So, you are actually going to isolate the DNA and RNA then you are going to check the quality of the molecule by running it on to the agarose gel. So, in the case of DNA you are going to run the regular agarose in the case of RNA you are going to run the denaturing agarose gel.

Once that part is done and you have cleared the QC part it is called as you know you are going to pass the quality control. Then you are going to have the primers and the probe design. So, the primers are designed that will bind to the region complementary to the target sequence. Probes are also designed which are highly specific for the target regions and the primers and probe play an important role in the quantification of the sample. So, once you are done with this you are going to set up the reactions and you are going to prepare the PCR reactions that will contain the cDNA template.

Some cases if you are isolating the DNA then you are going to have the DNA then you are going to put the forward and the reverse primers, you are going to put the fluorescent probes or the dye then you are going to put the DNA polymerase, reverse transcriptase, buffer solutions and dNTPs and then you are going to set up the amplification cycles. So, amplification cycle more or less remain the same as what we have discussed for the regular PCR that you are going to have the denaturations, then you are going to have annealing, then you are going to have extensions and so on and this will continue after the extension it will again go for denaturation and so on. So, exactly the same as what we have discussed for the traditional PCR. And then it is going to have the source and monitoring.

So, you are going to have the real time PCR. So, that is you do not need any kind of steps to be done that will eventually be done by the machine and it is going to measure the intensity of

the cyber green method or the intensity of the fluorescent dye from the Taqman probe method and then you are going to have the threshold detections. So, the predefined threshold level of fluorescent is set and that is actually been done by the experience ok. So, you will know that what will be the threshold values for this particular gene product or so. The cycle at which the fluorescent surpasses the threshold level is recorded as the cycle threshold or the CT values and the CT value is inversely proportional to the initial concentration of the sample which means if the CT values are low, then your concentration of molecule is concentration of molecule is high is very high because then that is how you are reaching to the CT value at very very low cycle number. Then you also have the 6 number that is the data analysis.

So, there are softwares available which detect the fluorescent signal and generate a standard curve which used as a reference which used as a reference DNA RNA concentration which allow the determine the initial concentration of the sample. And now once you are done with all these you can actually going to get the fluorescent data right the pattern how you are going to your how the fluorescent signal is moving into this and then you can actually be able to do the data analysis. So analysis of real time PCR. So in real time PCR the most common graph method is the amplification plot which usually represent the accumulation of PCR product over the course of the action. So this is what it is you are going to get of in on this side on the x axis you are going to have the cycle on the y axis you are going to have the fluorescence and it is going to have the multiple phases like the initial phase you are going to have the baseline then you are going to have the threshold values then you are going to have the plateau.

So in the x axis represent the PCR cycle or the time each PCR cycle involves the denaturation annealing and extensions y axis represent the fluorescence cycle detection by the instruments right and then you are going to have the amplification curve. So each sample or target gene is represented by the specific curve on the graph these curves show how the fluorescence signal increases over cycle a steeper slope indicate a higher amount of target DNA. So this pattern of this slope is also going to give you an idea what could be the amount of DNA present right because if the pattern is sharp if it is slowly moving then you are actually having the very low concentration of the target DNA. If it is moving very fast or steep then basically it is you know your amount of DNA is very high. So amplification curve has three phases the initial phase then you are going to have exponential phase and you are going to have the plateau.

So in the initial phase it represent the early phase of the cycle PCR cycle not enough amplification of cDNA to produce a significant amount of flow the cycle just started with denaturation of double standard DNA to separate the single standard DNA. So in the initial phase it is just going to prepare the machine or prepare the system for starting of amplifications. Then in the exponential phase there will be a you know there will be of amplifications and every after every cycle it is going to be double right. So in the exponential phase the flow sense signal increases with each cycle and it is going to be double after every cycle. So from it forms a well defined sigmoidal curve which indicates that the PCR reaction is proceeding efficiently.

Then during this phase the CT value is typically at the middle of the phase which is used to determine or quantify the initial concentration of the target DNA. The slope should be steeper

which tells the efficiency of the reaction while the linearity of the slope represent the doubling of the amount of DNA. Then you have the plateau phase at this phase the system is going to get saturated and you are going to get the plateau. So the plateau represent the point where the majority of the target DNA has been amplified. The curve loses its linearity since it is not the suitable for precise quantification of DNA.

And then you have the threshold line typically set the level where above the threshold baseline with flow sense the cycle at which the each curve crosses the threshold is called as the cycle quantification or the cycle threshold. The lower the QHC value higher the initial amount of target DNA. As it is known that CT value is inversely proportional to the initial concentration of graphs from the graph it can be concluded that the CT1 is smaller than the CT2. That means the DNA concentration in a small sample 1 is bigger than the DNA concentration in the sample 2. And that is what you are going to do the analysis and based on that analysis only you are going to say which gene or which gene product or which sample has more amount of DNA and less amount of or less amount of DNA.

So you can actually be able to compare the two samples. So this is all about the theoretical aspects of the real time PCR what we are going to do and what we have actually understood that what are the different steps you can be able to perform what are the requirements and how you can be able to perform. Now if you want to do the real time PCR in your laboratory you also need experimental feedback. You also require the practical demos so that you can be able to understand how it can be performed. So in the if you see the steps first step is how you are going to prepare the sample right. So in the first step you are going to denature the cells and you are actually going to prepare the RNA right.

And we have prepared a small demo clip so that you can be able to get acquainted how you can be able to isolate the RNA. Remember that in the past we have discussed about the RNA isolations whether it is using the quality affinity chromatography or whether we have used the other methods. So we have prepared a small demo clip and it is actually going to explain you how you can be able to prepare the RNA from the sample. Hello everyone in this video I will be discussing how to isolate RNA from the clinical samples.

So first step is to add the trizol in the samples. So this trizol that I have already added in it this trizol contains 40 percent of phenol, guanidine thiocyanate, ammonium thiocyanate and sodium acetate buffer. So after the addition of trizol we will add 200 microliter of chloroform in it. So I have added the chloroform in it then I will gently mix the solution. So until it turns milky. So after that I will centrifuge it at 4 degree Celsius for 15 minutes at 13000 rpm.

Now the centrifuge step is done we will collect this upper transparent layer from it and without touching the middle interface layer and transfer it to a new tube. Then we will add equal volume of isopropanol into this. Then gently mix the tube by simply inverting it. Now we just incubate this tube at the room temperature for 15 to 20 minutes. Now the incubation is done we will just centrifuge it for 15 minutes at 4 degree Celsius at 13000 rpm.

Now after the centrifugation we will just discard the supernatant and wash the pellet with 70 percent of ethanol. As the RNA quantity is very low we cannot see this pellet. So we will just blindly add the 70 percent ethanol.

Just try to detach it from the bottom. Just mix it two three times. Now after the mixing we will again centrifuge it for 5 minutes at 8000 rpm at 4 degree Celsius. Now we will just discard the supernatant and let the RNA pellet dry out and let the ethanol to evaporate for 5 to 10 minutes. Now as you can see this ethanol is completely evaporated. Now we will just re-suspend the RNA pellet into the QPSD water.

Here I am just re-suspending it in 20 micrometer. I am mixing it with a pipette so that all this RNA get diluted in this. Now we just incubate this RNA samples at 60 degree Celsius for 10 minutes. Now after this heating we just quick chills the RNA at ice for 5 minutes. Now the RNA isolation step is done. We can just quantify the RNA how much RNA we have just isolated and then we can just use for this like for cDNA preparation.

So now you have understood how you can be able to isolate the RNA from your samples. Once you isolated the RNA you are actually going to set up the reactions. And we are not getting into the detail of primer and probe designing because that we have already discussed when we were discussing about the traditional PCR. So primer designing is exactly the same but once you have isolated the RNA your second part would be that it gets converted into cDNA because you are not going to put the RNA into the reactions you are going to put the cDNA into reaction and that will be your second part. So that also we have prepared a small demo clip where the students are going to show you how you can be able to convert the RNA into the cDNA and you are going to set up the reactions.

Now as the RNA isolation is done in this video I will be showing you how to prepare the cDNA from the RNA. Now for that we need a I have prepared a I will be preparing a 10 microlitre of reaction and in 10 microlitre I will be adding 1 microlitre of random primers to prepare the cDNA. This is a random hexa what I will be using for this. For the cDNA preparation everything should be at a it should be after the primer I will be adding 2 microlitre. I have 10x reverse transcriptase buffer so for the 20 microlitre of reaction I will be adding 2 microlitre.

Now 5 microlitre is water. This is the RNA so I am using 3.2 microlitre of RNA because I have quantified the RNA that we have isolated and for 1 nanogram of RNA we will need 3.5 microlitre of RNA. Now the reaction which is done we will just set it up in the PCR. So, we have 2 standard PCR cycles first is 10 minutes for at 25 degree Celsius next is incubation or extension time for 2 hours at 37 degree Celsius and finally, we will just inactivate the whole reaction at 85 degree for 5 minutes. We just confirm the volume all the confirm the steps and keep the tube in the PCR.

Now we will try the block and we just start the reaction. Now as you can see it will take 2 hours and 15 minutes to complete the reaction and after that our cDNA is prepared. Now once you have prepared the cDNA you are going to set up the reactions. So, you are going to put the cDNA into the reactions right you are going to set up the RT-PCR reactions you are going to set the

amplification cycles you are going to see how the fluorescent monitoring is working or not and you are also going to set up the threshold detection right and these steps also we have prepared a demo clip so that you can be able to get familiarized with the how you can be able to perform the real-time PCRs and how you can be able to set up the reactions once you have the cDNA. In this video I will be showing you how to do the real-time PCR of just of the cDNA that we have recently prepared. So, my targets will be GAPDH and some virus primers like in this case I am using NDV and N primer.

So, for all the reactions I will be doing in the triplicates. So, as you can see these are the tubes for the real-time PCR and I am using 3 tubes 3 sets for a single sample. So, this is the sample and I will be amplifying with the cyber green. So, for a 10 microliter of reaction I will be requiring 5 microliter of cyber green. So, I am going to add 5 microliter in each tube. So, I am going to add 5 microliter of cyber green, 1 microliter of PCR sorry primer, each of the tube.

This is GAPDH, this is a virus primer. So, I am going to compare this amplification of viral RNA with the internal control of GAPDH. So, this is the sample, 1 microliter of primer, I will add 3 microliter of water to make the old reaction up to 9 microliter. Now, I have added 9 microliter of buffer which contains cyber green, primer and water. Now, the most crucial step is to add the cDNA which is 1 microliter.

So, in 1 microliter of cDNA I will be adding 5 nanogram of cDNA. So, to by adding equal amount of cDNA which is 1 nanogram, 5 nanograms sorry into each well I will be ensuring we have equal amount of RNA. So, that we can compare if we can compare the virus modulation or virus replication inside the samples. Now, the reaction is complete. So, I will just close the lid.

The reaction is setup, we will just put the these tubes in the real time machine. Now, we just setup the reaction of real time. This is the point studio real time machine and we will just setup complete run. Now, we will just fill up the data. So, we have just done the experiment test 1, this is the block type. We are calculating comparative C value, CT value and we have used cyber green for the amplification.

Just click next. This is the this is the PCR cycle. The total volume will be 10 microliter as you can see. This is the cold stage 15 degree Celsius for 2 minutes, then 95 degree Celsius for 10 minutes step 2 and final PCR stage or amplification stage it will be at 95 degree Celsius for 15 seconds and 60 degree Celsius for 1 minute and this complete cycle will be done for the 40 times. The final step is melting curve which is 95 degree Celsius for 15 second and just 60 degree Celsius for 1 minute and final is 95 degree Celsius for 1 second.

This is the dissociation step. So, this is the complete run complete cycle. We just click next. We do the advance setup. As you can see as I have already discussed, I have two different targets.

First is gap DH and next is NDV NDV N and we have only one sample. So, this is test sample. Test sample we just tell the machine or tell the software that which well we have used. So, I have

placed my tubes in the sixth lane. So, all these things are these six things are test sample.

First three are for gap DH and next is the NDV. We just hit next button. We just see if there is something in the machine.

It will show here the run number. So, this is the run number. Just click this. I will save here. Now, the reaction is started. So, it will take around 96 minutes to complete. So, this is the amplification plot. This is the number of cycles that we have put from as you can see this 0 to 40 cycles and this is the reaction threshold.

From here we will see this sample are getting targeted hits or not. So, these are the samples. Now, it will take 96 minutes to complete. Now, this is it for the real-time PCR. So, after this 96 minutes, we will just get the CT values. For that, after that we can just compare the gap DH and the NDV and CT values and see if there is an amplification of NDV and gene or not.

That is it for the review. Now, we have what we have discussed so far. We have discussed about how you can be able to prepare the RNA. Once you prepare the RNA, how you can be able to generate the cDNA from the RNA and then how from cDNA how you are going to take the cDNA and put it into the RT-PCR reactions, how you are going to set up the RT-PCR reactions, how you are going to set up the amplification cycles, how you are going to monitor the fluorescence and threshold. Now, once you are done with the threshold detections, you are going to get the data actually and you are going to get the fluorescent data. Now, at the last step, you are actually going to do the data analysis and data analysis so that you can be able to calculate the CT values.

You can be able to know what is the number of cycles in which you have crossed the CT values and so on. So, that also we have prepared a small demo clip which will actually going to explain where the students is actually going to explain how you can be able to compare the expression level of the two different genes from a particular two different samples. So, in this analysis, they have taken the GAPDH as the housekeeping genes and it is not going to change from one sample to another sample. So, that is actually going to be used for the background corrections or for equalizing the two samples and then they have also tested the other gene for experimental purposes. So, this is the analysis part. So, as I said at the end of the PCR reaction, we will get a CT value and then we have to compare how the expression of that gene is modulating due to the virus infection.

So, you can see I have already arranged the data. These are my samples, mock, these are treated or we can say virus infected and these are the repeat samples for second gene. So, one gene I am having GAPDH as a normalizing control and second I am having a p53 gene. So, in this case, we want to see how this p53 gene is getting modulated due to virus infection. So, suppose if I am having a control sample and the second one I am infecting the cells with the virus, so some proteins will go down regulated, some protein will go up regulated.

So, we want to see what is exactly happening to this p53 at the mRNA level. So, these are the

CT value and the method that we are going to follow to calculate the full change is this one. Analysis of relative gene, the method name is 2 raised to the power minus delta delta CT method. So, you can read this paper about the analysis. This paper is having almost 160,000 citations.

So, you can see, you can take a photo of that paper as well if you want or you can write the title. So, this is the analysis of relative gene expression data using real-time quantitative PCR and that this method. So, in this method to calculate the full change, we need to calculate this $2^{\Delta\Delta CT}$. So, in a minute, I will tell what exactly this is.

So, as I said, we have a CT value. So, you can see the CT value is 19.7, 19.365, 20.459. So, in all the three cases, this is almost similar. So, we will calculate a mean CT value, average of these three values. So, this will copy and we will calculate the CT value for each for all the cases.

Now, we have a mean CT value, say 19.8, this 25.8, then we have to calculate a delta CT. Now, this is mean CT value. Now, we will calculate delta CT, difference of CT value. So, which difference? So, in a single sample, we have two genes GAPDH and a P53 gene and we have two CT values. So, we will calculate the difference between these two.

So, for every difference, we will have a two value, this one and this one. So, first I will use the value of control, this minus, this value. Now, we have a delta CT. Similarly, we will calculate the delta CT for the treated samples.

So, this is delta CT. Now, we will calculate delta delta CT. Delta delta CT is now we are using, as I said, we are using GAPDH as a normalizing control. So, suppose I am giving an infection and some of the protein, as I said, is going down regulated. So, the cells will probably die. So, the number of GAPDH mRNA in all the cells will be different.

So, we want to normalize the expression of all the proteins using a GAPDH. So, see, the GAPDH is, in this case, it is 19.8, but in the infected one, it is 25.

8. So, you can guess in which of the cells the GAPDH expression is more. Obviously, 19.8, but we want to normalize it. So, if we are getting the value of GAPDH in this case, P53 is 25. But in the infected one, the GAPDH is 25.8 and this is 29.

5. So, we can't say if the virus is getting, due to virus infection, the value of P53 is down regulating or up regulating. Can you guess it? No, we can't say for sure. Because in both the cases, the GAPDH is different.

Suppose we are getting the same GAPDH here. So, in both the cases, we are getting 19.8 value for GAPDH. And then we are having the P53 at 25 and then 29. Then we can clearly say because of this virus infection, the P53 is going up or down? Not up regulated, down regulated because the CT value is increasing.

So, it is down regulating. But in this case, we are not having a similar GAPDH. It is slightly

increased. So, the GAPDH is also less. So, we need to normalize this value to this value.

So, we can compare the expression of P53. So, this is normalization. First is calculating the delta CT. Then we will calculate the delta delta CT. Delta delta CT will normalize by subtracting this value with this value only. So, I will repeat it again. Delta CT is the difference between the mean CT value of two genes from the same sample.

Difference between the mean CT of two genes from the same sample. So, we calculated the difference between here and this. And this delta delta CT is normalization basically. So, we are going to normalize this value with this only. So, that's why I am subtracting the same value.

But in this case, we are not going to normalize. We are going to simply use subtracting this value with the test one. So, here we have minus 1.6. Now, the fold change is as I said the method is 2 raised to the power delta delta CT.

Now, as you can see we got our delta delta CT. So, we will raise the power. 2 raised to the power minus of delta delta CT. So, you can see fold change is 1 which means we have successfully normalized the GAPDH value because it's coming 1. But we have to calculate the fold change for this. So, you can see it is coming 3. So, which means that if we are normalizing the value in the mock treated, the expression that we have normalized is 1.

But in the treated one, it is 3 times. So, it is up regulating the gene, the p53 gene. But initially when we saw the value, we could see that 25 or 29, we could see that the value was getting increased. Because of this gap, we are able to see this is not decreasing, the expression is higher as compared to this one. So, this is how you plot a graph. So, in the mock control, you will have value 1 and in the treated one, you can see this is 3.

So, suppose if I am changing the value and decreasing the value, suppose I am 22. So, this is 21 and this is 22.5. So, you can see this is how this value is increasing. Now, this is 655 times of the one.

And say if I am decreasing this value, say 35, this is also 35 and this is 35, 24.5. So, you can see now this value is 0.08, which means the virus down regulated the gene. So, do you have any doubt? So, this is the analysis part. Now, as Sir was mentioning about the COVID-19 or say like you generally went to, suppose you went to a doctor and you gave your sample for the testing, like if you are COVID positive or not. So, what will you do? You will collect the samples and will run the PCR and at the end, you will get a CT value.

Now, as I said, we are setting up 40 reactions. So, out of 40 reactions, at which cycle we are getting the blood pressure? In this case, 25. In this case, 90. So, suppose if you are COVID positive, then you will get a CT value lesser than 40. So, suppose you are getting 40 value of 35 and my friend, suppose I am getting a value of 35 and my friend is getting a value of 29. So, the doctor will say that guy is more infected with the virus because the CT value is less and I am less infected.

So, this is just telling by the CT value if you are infected or not. But if you want to see how the gene is getting modulated to the infection, so we can do this. At the end, we will get a fold change and you can plot and you can just simply publish your data, whatever you are getting at the end. So, this is experiment number one. So, we have to, as I said, this is a quantitative data.

So, we have to repeat the same experiment with biological replicates and at the end, we will get a three-fold change. Suppose in this case, as I said, we are getting fold change of 3.

08. In the next experiment, we will get something around 3. In the next one, suppose we are getting 2.0. So, we will take average of all the three experiments and then we can do the statistical analysis and then we can see our data. We can say that our data is significant and this is the final value.

So, this is the analysis part. So, if you have any doubt, you can ask now. Housekeeping changes. So, if you don't have a doubt, we can end this session. Thank you. So, this is the analysis part. So, as I mentioned in my earlier session that at the end of the PCR action, we will get a CT value.

So, these are the CT value that we got and this is the MOF control and the CT values are there. And this is the virus infected sample. So, I am using two genes, a GAB-DH as a normalizing control and my TANF gene, a p53 gene. So, in the analysis, we are going to see after infection what is happening to the p53 gene expression. It is up regulated or down regulated due to the virus infection. The method that we are going to follow is this method, analysis of relating gene expression data using 2 raised to the power minus delta delta CT method.

So, you can take a picture if you want of this paper and read it thoroughly. So, at the end of this analysis, we will be calculating this value. So, I will be explaining this what exactly delta delta CT is. So, now we have a CT value.

So, you all know what the CT value is already. So, you can see that we have three values. So, for that we need to calculate the delta delta CT. So, first we will calculate the mean CT value of all the values.

So, first we will calculate the average. This is the average. I will copy, cut, copy and paste it here. This is the average of p53 gene. Then we will calculate the mean CT value. After that we will calculate the delta CT value. Delta CT value means difference of CT value. So, if you want to see the expression of p53 gene, we want to, first we have to normalize the expression.

Because you can see that in the first case our CT value is 19 in GAPGAT and in the second one it is 25. So, can you guess like in which of the cases the expression of GAPGAT is more? In the first one. So, obviously this is very much high. So, the expression is less because of virus infection. So, if you want to see the expression of p53 gene, what exactly is happening with the p53 gene, you have to normalize it with a caret here.

The word normalization means, can you guess like this is the CT value is 25. Around 25. In this case, the CT value is 35. So, these are the values. So, can you guess which of the cases the expression of p53 gene is more? In the first one. Now, you are guessing just because of these values. But normalization means like first we have to normalize with a single gene.

In this case, the GAPGAT is different. But if the GAPGAT is 19 or 20, in that case we can say the p53 expression is same. But due to the difference between the CT value in these two things, we can't say like the expression is more in which case. So, we have to normalize or we have to cancel out the difference between these two values of GAPGH of both the samples. So, we have mean CT value as we have calculated, then we have to calculate the delta CT value.

Delta CT value as I said is a difference between the CT. Which difference? The 19 in this 25. So, we will calculate the difference. So, first I will take 19 minus this value. Now, this is delta CT value. Similarly, we will calculate a second CT delta CT for infected samples.

Now, this is delta CT. Then we have to calculate delta delta CT which means difference of delta CT value. Now, how many delta CT values we have? Two. So, we have to normalize it this one. So, first as I said we will be normalizing the value of expression. So, what we have to do? We have to subtract these values with this one only because this is control sample and we want to normalize it.

We want to see what exactly is happening in this case. So, we make the value of P53 expression as 1 in this case and then we will calculate the full chain in this case.

Are you getting my point or you are simply saying yes? Repeat. Yes. Okay. So, first we calculated the CT value. Okay. Then after the CT value we calculated delta CT. Difference between the CT value. Now, you can write the definition of CT value delta CT.

Delta CT is difference between the CT value of two genes from the same sample. Difference between the CT value of two genes within the same sample. So, this is delta CT. Difference between this value of P53 and GABH from the MOB treated. And in this and other case, difference between the CT value of P53 with GABH is the treated one. Next, we have to calculate the delta delta CT.

So, as I said we are going to normalize the expression of P53 as 1. So, for that we need to subtract the same value with the same value only. This value with same value. So, the delta CT value will be 0. In the next case, now we are going to see the expression.

So, for that we need to normalize with the first one. So, we will subtract the value of delta CT with the MOB treated one. Okay. So, the delta CT is now you can write the definition of delta delta CT.

It is the difference of delta CT value between two samples. Okay. Between the two samples.

In this case, as we are normalizing, so we subtract it with the same value. But in this case, we are not normalizing, we are calculating the, we are checking the regulations. So, that's why we are subtracting with the control one.

Now, the fold change. So, as you can see from the paper, this is 2 raised to the power minus delta delta CT. Now that we got this delta delta CT, we will calculate the fold change.

What the fold change is? 2 raised to the power minus of this value. Okay. But the fold change is 1. Why we are getting the fold change as 1? Because we normalized it to 1. Okay. So, if we calculate the fold change for treated one, we are getting 3.408. So, in the initially, initially you were saying the expression is more in this case in the control one.

But after analysis, what we got? The expression is more in the infected one. So, for graph, the mock treated one will be 1, the treatment will be 3.08. So, what we got at the end, the virus infection is increasing the expression of p53 by three fold. Okay. So, this is the expression. Now, as I mentioned earlier, like we generally prefer the real-time detection method.

Like if you remember, you can recall this thing, like you went to a doctor and you got your real-time PCR if you are COVID positive or not. So, what he did, what he or she did, he took your samples from your nose and they did the real-time PCR and at the end, they got the CT value. Now, the CT value that we got was ranging from, was lying between from one to 40 because we are running the reaction for 40 cycles. So, suppose I am getting the CT value of 35 and my friend is getting a value of 32 or 30.

So, which of us are having more infection? 32. Obviously, but if you want to analyze the expression of a particular gene in both the cases, then we have to use the GANs. And this is the analysis method and this is how you calculate the whole gene and you can plot a graph and then publish it. This is a replicate first. We have to again do the same experiment, biological replicate and at the end, we will get a similar value.

Then we can take the average and see how significant our data is, then you can publish it. Okay. Do you have any doubt? We are isolating the mRNA from the cells, then we are reverse transcribing it to the cDNA and we are using that cDNA as a template in this reaction. We are not isolating it for the, no we are not removing anything.

This is one simple, one step, yes. Okay. So, after the end of the analysis, you could have understood that how you can be able to perform the real-time PCR. You are going to prepare the samples preparations, you are going to isolate the RNA, you are going to prepare the cDNA, then you are going to put the cDNA into the reaction mixtures, you are going to put the, you are going to set up the amplification cycles. All these depends on the type of products you want to develop and then from that you are going to get the data and then that data is actually going to be analyzed thoroughly. So, that you can be able to know which protein or which gene is being amplified or it is actually, you know, showing the lower expression. So, this is all about the real-time PCR.

We have discussed, what we have discussed, we have discussed about the basic principle, why we are using the real-time PCR, what is the advantage of using the real-time PCR and so on. And at the end, we have also discussed about the experimental aspects. So, we have prepared, we have shown you the couple of demo videos, how you can be able to perform the real-time PCR in your laboratory. So, with this, I would like to conclude my lecture here. In our subsequent lecture, we are going to discuss some more aspects of molecular biology. Thank you. .

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 10

Molecular Cloning

Lecture-40 Cloning (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biotransition and Bioengineering IIT Guwahati. So, in today's lecture, we are going to discuss about the enzyme production. So, enzyme production as the name suggests is where you are actually going to use the where you are going to prepare the large amount of enzyme and we are there are two approaches one can actually be able to use. First approach is that one can actually be able to purify the enzyme purify the enzyme from the rich sources. For example, if I have to purify the enzyme of the glycolysis for example, if I have to purify the enzyme of hexokinase, then what I can do is I can just isolate the cytosol of a cell and then I can use the different types of conventional chromatography techniques and I can be able to get the pure enzyme.

Same is true for many other enzymes like I can use for lysosomes, I can use the lipase, I can use any other method. So, in this particular approach, you have to go and first identify the rich source. For example, if I have to purify the lysosome, lysosome is very heavily available and in many of the biological fluids. So, I have to identify those biological fluids and then only I can be able to use the different chromatography techniques to get the pure enzyme.

The second approach, which is more straightforward and which actually gives you the, you know, the enzyme with lot of ease is that you can actually be able to identify. So, enzymes are made up of the proteins and the proteins are being synthesized from the RNA, which is actually going to be coded from the DNA. And this DNA is nothing which is also called as G. So, for example, if I have to work or I have to use an enzyme for my experimental applications or if I want an enzyme for structural elucidations or if I want an enzyme for the industrial applications, then I can just first do the enzyme, I should identify the corresponding gene. And once I identify the corresponding gene, then what I can do is I can just clone this gene into a suitable vector.

And once I have the gene into the suitable vector, I can do the over expression and I can be able to purify this protein with the help of the conventional chromatography and as well as the specialized chromatography. So, once I purify the protein, I am going to get the pure enzyme and that pure enzyme I can be able to use for many applications. For example, if I

need an enzyme for studying the DNA recombination or if I need an enzyme for running the transcription, translations, replications, all that, so that all I can be able to prepare in large quantities. And so that you can be able to do because there is no, once you have done that into a, once you have cloned the gene into a suitable vector, then you can actually be able to, you know, transform that into a suitable host and that is how you can be able to make the proteins in large quantities. So, this is what we are going to discuss in this particular module and in the subsequent module as well, that how you can be able to use this kind of approach to produce a large quantity of enzyme because this enzyme actually be able to use for many applications such as we can use it for the structure elucidations, we can use this for catalyzing the reactions and we can also use this for the industrial applications.

Now, if I have to go with the approach number 2 or if I have to go with the approach number 1, we have to have a very clear understanding because the, see with what is the drawback of the approach 1 is that first of all you have to go with and identify a rich source, the second is you have to purify this protein with utilization of the different additional chromatography technique and even then the production is going to be very, very limited because it depends on the amount of enzyme which is present in this particular rich source. Whereas in the approach number 2, once you have identified the gene, you can actually be able to overexpress that into a very large quantity and also since you have the gene, you can be able to do the fine tuning. For example, what is mean by the fine tuning is that suppose I am working with an enzyme and I want to identify its structure with a larger objective is that I want to study how this enzyme is interacting with the substrate and how it is actually forming the complexes with the substrate and product and so on so that I can be able to get the inner insight into the enzyme catalyzed reactions. In that case, I might actually require the enzyme alone and I might also require the mutants, mutant of those enzymes so that I can actually be able to study the enzyme, how the enzyme is interacting with the substrate and then I also can test the mutants. The mutant approach is not going to work when you are going to work with the approach number 1, but the mutant approach will work in the approach number 2 because you have the gene and you know at what point you are actually trying to make the mutations and that you can actually be able to insert into this particular gene and that is how you can be able to change the overproduced enzyme.

So this is one of the major advantage of going with the approach number 2. So in the approach number 2, what you are going to do is you are going to first identify the gene. So the first step is that you are actually going to identify the gene and that you are going to do by screening the different types of clones that we are anyway going to discuss in detail. So that you require a screen, right? Because you have to screen the gene in a pool of genes so that you can be able to identify the gene which is responsible for the production of the enzyme X for example, okay? So once you identify the gene, then you are actually going to clone that gene into a suitable vector. This cloning also has the multiple steps.

In the step number 1, you are going to amplify this gene with the side-specific, with the help of the side-specific primers and these side-specific primer are actually going to have

the restriction enzyme. So once you have the restriction enzyme, they will be able to, you can be able to cleave that amplify gene and that is how it is actually going to produce the sticky ends. And these sticky ends are then going to be, you know, be facilitate the entry of this particular gene into a suitable vector and that is how you are actually going to produce the clone. So in the step number 2, you are going to do the restriction digestion so that the sticky ends are going to be generated and then you are actually going to clone that into a suitable vector. Once you clone it into a vector, then you are also going to do a screening of these clones so that you can be able to select the suitable clone.

So you are going to use the suitable clone. Now once you got the suitable clone, then you are going to use that suitable clone for overexpression. So in the step number 3, you are going to do the overexpression. In the overexpression, the first step is that you are going to choose the host cells. Host cell means the cell where you are actually going to make the proteins.

It could be very simple such as the bacteria or E. coli. It could be as complicated as the mammalian cells. So depending on the host cells, it can be yeast, it could be insect cell lines and all that. So these are the multiple options what you have at this stage when you want to use the host cells.

And choosing the host cell is very crucial because that is going to decide the final production. Then once you have selected the host, you are going to do the transformation of this particular clone what you have produced in the step number 2. So this is the step number 2. This is the step number 3. And once you do the transformation, after that again you have to do a screening.

The third step, you are going to do a screening to select the suitable transformed clone. And then you are going to do overexpression. For overexpression also you are going to have the different strategies depending on what host you are going to use. And then ultimately you are going to do the analysis of the overexpression which means you are going to see whether the protein is been, protein or the enzyme is being produced or not.

Okay. So once you are confirmed that the enzyme is getting produced, then you are actually going to make the large cultures, you are going to produce large amount of host cells and that is how you are going to have the large amount of proteins. But these proteins are actually also going to have the host proteins. So when you overexpress you are going to have two proteins, the enzyme X, which is going to be in an overstressed state and then you also going to have the host proteins. This means you are going to get a mixture of the enzyme X and the host protein after the overexpression. So then you are going to enter into the step number four where you are going to do the purification so that you can be able to separate out the enzyme X from the host protein and that is how you are going to get the pure enzyme X.

Now when you want to do the purification, you also again going to do the multiple step. In the step number one, you are going to do the cell disruption. In the step number two, you are going to do or you are going to perform the different chromatography techniques such as ion exchange chromatography, hydrophobic interaction chromatography, gel filtration chromatography or sometime you might also do the affinity chromatography. Once you are done with the chromatography, then you are going to do the estimation of the or the yield of your purification or the yield of what amount of enzyme X you have produced and then you are also going to do the further analysis for seeing how much is the purification. So, all these you are going to do.

So, all these four steps are required if you want to produce the enzyme under the approach number two. So, let us start discussing about these approaches and these steps one after the other. So, today we are going to discuss about how you can be able to identify the gene of corresponding to the protein X or enzyme X. Now there are different approaches for isolation of the gene of interest like for example, in this case, we want to identify the gene which is responsible for the enzyme X. So, what are the things you can actually be able to use as the starting material? So, if I have the enzyme X and if I so, it depends on what kind of phenomena we what kind of you know information we have.

So, for example, if the gene sequence is not known, for example, this is one condition where the gene sequence is not known. If the gene sequence of this enzyme X is not known, then you are going to use these two approaches and if the gene sequence is known, then you are going to use this approach. So, in that if the gene sequence is not known, you are actually going to know either two information one, you are going to have a gene fragment, which you are sure that it is actually expressing a part of enzyme X or you are actually going to have the messenger RNA or you are actually going to have the active enzyme, which is going to be active. So, if you have the gene fragment, you can be able to go and look at into the genome of their that particular organisms. So whether you want to use the genome or whether you want to use the expression profiling, in both of these cases, you might have to prepare a genomic library.

So, in this case, when you want to look for the genome, you have to prepare a genomic library. So, genomic library is nothing but a collection of the clones, collection of clones corresponding to different genes. Similarly, so all the genes, whether these genes are expressing and giving you the protein or not, they are actually going to give you all genes. Whereas when you are going to go for the expression analysis, you say that you have the active enzyme but you do not know which gene actually, then you can actually be go with the transcriptome or you are actually going to go with the cDNA library. So, when you are going to go with the cDNA library, cDNA library is actually a collection of the X gene expressing in a cell.

So, all the genes, so this is actually a expression library, this is the genomic library. So, it is actually going to take care of the all the genes, which means genomic library will contain

the cDNA library, but cDNA library will not going to contain the genomic library. This means the seed whatever the genes you will see under the cDNA library will also be present in the genomic library, but whatever the genes you will see under the genomic library may or may not present in the in the cDNA library. Once you got the library site, just like as you have the library in your college or you know in your institution, and you want to see that, I want to today I want to study a book which is about the protein biochemistry or protein structure determinations, then what you are going to do, you are going to take these keywords and you are going to put it into the software. And what the software is going to say, software is actually going to go through with the your institutional library and it will actually going to tell you, these are the books which are available which is related to the protein structure determinations.

So, then you are actually going to screen out the suitable book actually. Similarly, these are the collection of the clones, these are the collections of the genes, one is for the whole genome, the other one is only for the expressions of expression status of the protein, expression status of that particular organism. So, then the second step is, so in the first step, you are actually going to prepare the genomic library or the cDNA library. In the step two, you are going to screen the library with the help of the gene fragment or the active enzyme what you have with you and then you are actually going to use the screening. In the step three, you are going to take out or you are going to isolate these gene fragments and then you are actually going to clone it into a suitable vector.

And that is how you are actually going to get the final clone. So, this is all about the approach when you do not have, when you do not know the genomic sequence, which means the pre genomic era, before the pre genomic era, people were actually going with this approach where they are actually only having the, they have isolated an enzyme with the help of the traditional approach, but they do not know gene sequence. So then they actually go with these kind of approaches. After the post genomic era, when the people actually know the genomic sequences, then you are very, because you know the genomic sequences, you can actually be able to use a technique which is called as polymerase chain reactions. And that will, it is actually going to give you the amplified gene product and that you can be able to clone it into a suitable vector.

And that actually can you can be used further into the over expression and purification steps. So, first we will discuss about the approach number one where the genomic sequence is not known. And then we are going to discuss about the gene, the approach where the gene sequence is known. Now the first approach is when how to prepare the genomic library. So preparation of genomic library is a multi step process.

In the step one, you are going to isolate the genomic DNA of the host or of the, you know the, of the organism from which you are interested to isolate the enzyme. Then you are going to do the restriction digestion of the genomic DNA or in, or in general what you are going to do is you are going to make the fragments. Then the step three, you are going to

purify these fragments so that you can be able to use that. And then you are going to ligate these fragments into a suitable vector. So in this case, we have taken an example of VAC, YAC vector.

And then in the step five, all these are going to transform into the yeast so that you are going to get a library of clones. And once you got the library of clones, that is actually going to give you a genomic library which is actually going to represent the, all the genes which is present in this particular organisms. So how you are going to do this? You have, what you are going to do is you are going to, first what you are going to do is you are going to isolate the genomic DNA or the human genome, because we are talking about the preparation of the genomic library from the human. And then you are going to digest that human genome. So all the chromosomes are going to be digested with the help of the excision enzymes.

And that is actually going to give you the different fragments, all these fragments, which actually contains one or more genes in each fragment. Ideally, you are going to produce the fragment so that it represent at least one gene, it may have the additional fragments or the 50% fragments more from the other genes. And the same way you are actually going to take the YAC vectors. So YAC is the yeast artificial cloning vectors. And that actually you are also going to have the EcoR1 site.

So you are actually going to digest that with the EcoR1 and BamH1 and that is actually going to give you the linear fragments. And then what you are going to do is you are going to do the ligation. So once you go to ligation, all these fragments randomly and are going to be get ligated into the YAC vectors. And then you are going to do the yeast transformations and that is how you are going to get a clone. So this is one clone, this is second clone, this is third clone and so on.

So these are the different clones and all these clones are actually going to constitute the genomic library. So let us discuss about each and every step how you are going to first prepare or how you are going to first isolate the genomic DNA. So in the step one, you are going to do the isolation of the genomic DNA. So isolation of the genomic DNA is a multi-step process. In the step one, you are going to do the lysis of the cell with a detergent called lysis buffer.

With a detergent containing lysis buffer. So if the cells are cultured cells or the tissue, you are actually going to do the homogenization and that is actually going to give you the single cell. So once you do the, if it is like tissue, for example, if it is a liver or spleen and you want to prepare the genomic library, then you are going to use, first you are going to homogenize these tissues so that they will actually going to give you the single cell suspension and then these homogenized projects or the single cell is actually going to incubate with the lysis buffer so that it is actually going to lysis the cells. So you are going to get the lysate and this lysate is actually going to contain three components. One, it is actually going to contain the genomic DNA or gDNA. It is actually going to contain the protein and it is also going to

contain the lipid and so and it may also contain the RNA which is also present in the cell.

So these are the four material what you have and you are only looking for the genomic DNA. So what we are going to do is we are going to do the purification steps. So incubation of cell with the digestion buffer containing the proteinase K, SDS will release the genomic DNA from the DNA protein complex and this genomic DNA which you are going to get just after the lysis is actually going to be present as a complex with the protein and you are going to incubate that within a protease and that protease is actually going to chew up all the proteins and that is how your genomic DNA is going to be released. Then the step three, you are going to isolate the genomic DNA by the absolute alcohol precipitation. So we are going to precipitate the DNA but this DNA is also going to contain the contaminating agents like the proteins, lipids and RNA.

So it is actually going to not only going to contain the purified genomic DNA, it is going to contain the other molecules. So in that case what you are going to do is then you are going to do a purification of the genomic DNA with the phenol chloroform mixture and chloroform mixture has two phases. One is aqueous phase and other one is a organic phase. This is what you see right, it has an aqueous phase and organic phase. In this step the phenol denature the remaining protein and keep the protein in the organic phase.

So protein is actually going to be separated out and it will go into the organic phase. So the genomic DNA, the blushed colored phase what you see is actually going to be isolated again and it is actually going to be precipitated with the help of the absolute alcohol. Now, in the last step because it is a genomic DNA, so you also have to test that the genomic DNA is perfect. So you are going to analyze the genomic DNA on 0.8% agarose and a good preparation of genomic DNA give a intact band with a no visible strain.

So this is what you see, this is genomic DNA, this is the control where no genomic DNA and this is the genomic DNA. So what you see here is that we have the intact one band and that migrates very slowly because genomic DNA is very big. So it does not migrate very fast. So this is all about the genomic isolation of the genomic DNA and then once you have generated genomic DNA then the step two is the generation of the suitable size fragments.

Suitable size fragment can be done by the two methods. The first one is a restriction digestion method where genomic DNA can be digested with a frequent DNA cutting enzyme such as the EcoR1 or the BamH1 or SAW3A to generate the random sizes of genomic DNA. The criteria to choose the restriction enzyme or the pair of enzyme in such way such that the reasonable size DNA fragment will be generated. What is reasonable size is that it should contain at least one gene. It may have a flanking sequences that is not an issue but it should at least contain one gene in one particular fragment. So that it should not be the case that if you fragments are so big that it actually contains multiple genes because in that case it is actually going to fail the purpose of making a genomic DNA.

Ideally the genomic library means that you every clone which you are going to make should represent only the single gene or it should may have some flanking genes but only very fragments. It may have the fragments of other genes but at least it should have one gene. It should not have like four or five or six or more genes. So that is why you have to generate the fragment which are reliable and relatively big enough. It is likely that each and every genomic sequence is represented in the pool.

As size of the DNA fragment is large the complete genome will be presented in very few number of clones which means the clone number should be very high so that it should represent the all the genes which represent in that particular genome. So that is the step number the method number one where you can do the restriction digestion. In the step number two you can also do a mechanical sharing. So what you can do is just vertex the genomic DNA and when you vertex the genomic DNA it is actually going to break the genomic DNA into the multiple fragments. So genomic DNA can be fragmented using a mechanical sharing but restriction digestion method is more popular and more precise because here you are using an enzyme so enzymes are going enzyme mediated events are going to be more and more regular and more controlled.

Now for example how much would be the fragment size. So for example if an organism has a genome size of 2^{10} to power 2 into 10 to power 7 kB and an average size of fragment is 20 kB right. Then number of fragment would be n is equal to 10^7 divided by 20. So you divide this number by this number you will get this number. In reality this is the minimum number of clones to represent the fragment in the library where the actual number is much larger it almost be a 10 times the number what you have to put or what you have to get. So the probability of finding a particular genomic sequence in a random library of n independent clone is n is equal to $\ln n$ minus 1 divided by $\ln n$ minus 1 by n where n is the number of clones what you have in the genomic library p is the probability of finding a fragment into the genomic library and n is the size of the average fragment size.

So you can use this particular formula to calculate how many clones I should have to generate and that actually is going to increase your probability of finding the clone or finding your gene into a genomic library. So once you generated the suitable size fragment then it has to be ligated into the suitable vector. So in the step 3 you are going to use the cloning into the suitable vector. So depending on the size of the you know the size of the different organisms for example the mycoplasma bacteria or like the flowering plants you have to choose the suitable vector. So cloning into the suitable vector the suitable vector to prepare the genomic library can be selected based on the size of the fragment of the genomic DNA and carrying capacity of the vector.

Size of the average factor can be calculated and accordingly a suitable vector can be choose. In the case of the fragment generated by the restriction enzyme the vector can be digested with the same enzyme and put for ligation to get the clone. In the case of mechanical sharing mediated fragment generation putting these fragment needs additional

efforts. In one of the approach adapter molecule can be used to generate the sticky ends. Alternatively, endonucleases can be used to generate the sticky ends.

So these are some of the you know the vector carrying capacity of some of the vectors or the fragment carrying capacity of some of the vectors. For example the plasmids they can actually be able to carry the 15 MB phage lambda, cosmids, bacteriophage, back and the yeast. So depending on the size of the fragment what you are going to get from a particular genome you can be able to use the different vectors. Then you are in step 4 you are going to do the transformation to get the colonies. So the post ligation the clones are been transformed into a suitable host to get the colonies.

A suitable host can be bacterial stain or the yeast. Different methods of delivering a clone into the host is discussed in a future lecture. So when we are going to discuss about the over expression that time we will discuss about the transformation how the different methods can be used to deliver the DNA into the different host cells like the bacterial cells, yeast cells, mammalian cells and so on. So this is all about the genomic library approach where you have now prepared the genomic library. And now we will move on to the next approach where we are going to use the cDNA library. So that we are going to discuss in our subsequent lecture. Thank you.

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 10

Molecular Cloning

Lecture-41 Cloning (Part 2)

Hello everyone, this is Dr. Vishal Debeli from the Department of Biosciences and Biosciences and Biosciences. So, in this context, in the previous lecture, we have discussed about the approach which is related to the condition when the gene sequence is not known, right, which means these are the conditions when the pre-genomic era, the people were only knowing the property of the enzyme, but they were not very sure about the, what is the gene sequence. And in that case, either you will have the information about the genome of that particular organisms or you will actually going to have the expression of that particular gene. So, in the previous lecture, we discussed about the, how you can be able to identify the gene if the genome is known. So, in that context, what you have to do is you have to prepare a genomic library and genomic library is going to be a combination or the collection of the genes which are going to be cloned into the individual clone and it is going to represent the complete genome, which means the genomic library is also going to represent the cDNA library as well, okay.

Whereas, if the protein or the, you know that the protein is getting expressed and you have actually some tools which are, you know, going to be tell you that, okay, there is a, you know, expression construct or you have the antibodies or something, right, and you want to screen these, you know, the expression clones, then you have to, then you have to isolate the transcriptomes. And in that case, what you have to do is you have to first isolate the messenger RNA and then from the messenger RNA, you have to prepare the cDNA lab, cDNA and that cDNA, you have to clone into the suitable vector and that is actually going to give you the cDNA library. And once you have the cDNA library, you can be able to use that for identifying the clone of your interest. So, in today's lecture, we are going to discuss about the, how you can be able to prepare the cDNA library and then how, whether it is the genomic library or whether it is cDNA library, how you are going to screen the clones for identifying the clone of your interest and how you can be able to then isolate the particular fragment so that you can be able to clone that into the expression vector and that is why you can be able to produce the protein in the large quantity.

So, the construction of the genomic library, construction of the genomic library is a multi-step process where you are, in the step one, you are going to do the isolation of the messenger RNA. In the step two, you are going to prepare the complementary DNA or

cDNA in short and then the step three, you are going to clone that into a suitable vector system and in the step four, you are going to transform that into the suitable host so that you can be able to prepare the cDNA library. So, cDNA library is a collection of the clones which are going to show you the expression status of that particular cell. So, first thing what you have to do is you have to first break open the cells, you have to isolate the messenger RNA. So, you have to isolate the messenger RNA.

The step two, so first you are going to break open the cells. You are going to get the cell lysate and from the cell lysate, you are going to isolate the messenger RNA and we are going to discuss how you are going to do the messenger RNA and once you have the messenger RNA from the cell, so these are the total messenger RNA and then you are going to use these messenger RNA and with the help of the enzyme reverse transcriptase, you will be able to synthesize the DNA which is called as cDNA. cDNA means complementary DNA, right? So complementary DNA to all the messenger RNA what you have isolated and then once you have the complementary DNA, then you can insert that into the bacterial plasmids or other suitable vectors. So that is actually going to give you the different clones and then these clones are actually going to transform into the suitable host and that is how you are going to get the cDNA library. From the cDNA library, you are actually going to isolate the clones of your interest and that is how you are going to identify the DNA sequence and then you can be able to use that for subcloning that into the expression vector and that is how you are going to use that for protein production or the enzyme production.

So before getting into the step number one where we are going to talk about the isolation of the messenger RNA, we have to first understand the structure of the messenger RNA because that is actually going to lead to the approaches what you can actually be able to use to isolate the total messenger RNA. So if you see the structure of the messenger RNA, in the structure of the messenger RNA you have a 5 prime cap, which is actually the 5 prime cap, then you have the 5 prime UTRs, then you have the coding sequence. So this is the coding sequence which is actually going to be responsible for production of the protein because this is the region what is going to be transcribed, which is going to be translated by the ribosomal machinery and then you have the 3 prime UTRs, these are the, so 5 prime UTR and the 3 prime UTRs are the regulatory region and after this you are actually going to do the post translational modification. So one of the classical post translational modification is that you are going to add the poly A tail and this poly A tail is present in all the messenger RNA, the length could be different, some of the having the 200 messenger As, some can be having the 50 and so on. The amount of the As what are going to be present on the 3 prime end of the messenger RNA is going to decide the age or the age of the messenger RNA or the durability of this messenger RNA, this means it is actually going to decide the stability of this messenger RNA in the cytosol.

Now this is the region which actually one can use to identify and purify the messenger RNA. So what you can do is you can actually be able to design affinity column and that

affinity column is actually going to have a very high affinity for a poly A tail and that is how you can be able to use that to isolate all the messenger RNA what is present in the cytosol. So these are the different steps what you are going to follow for the isolation of the messenger RNA. So these are the this is the step 1 in the construction of the cDNA library. So first you have to isolate the messenger RNA so exploiting the structure of the messenger RNA you can actually be able to use the poly A tail and the poly A tail has a very high affinity for a poly T column because the A is having a very high affinity for T and A is having the complementary to the T and that is why you can use a poly T column and that poly T column is actually going to give you all the messenger RNA which having the poly A tail.

So what you are going to do is you are going to first culture the cells and then you are going to take those culture cells and you are going to do the lysis input it into the lysis buffer. If you are going to work with the tissue for example if you are going to work with the liver or spleen then in that case first you have to do the homogenization and the homogenization is actually going to break open the cells to a single cell suspension and single cell suspension is then going to be incubated with the lysis or the binding buffer and that is actually going to give you the lysate. Lysate means you are going to get the cell lysate and the cell lysate is going to have the two things one is you are going to have the messenger RNA is going to have the protein and it also going to have the other messenger other RNA species but as we said that messenger RNA is actually having the poly A tail so it is actually going to bind to the column. So what you are going to do is you are going to take the beads which are actually having the poly T linkers. So in these kind of beads what you have is you have a bead and on this bead you have a linker on which you are actually going to have the T nucleotides which are attached.

So these all these T nucleotides are actually going to have the affinity for a poly A tail what is present on the messenger RNA and that is how it is very specifically going to bind all the messenger RNA what is present in the cell lysate. So once it is once the binding step is over then you can wash the beads with a washing buffer and you can just discard the supernatant and you can collect the beads and then you are going to do the elution. So what you are going to do is you are going to do the elution with the help of either the poly A sequences or you can change the pH and other things. So once you do the elution you are going to have the two fractions you are going to have the bead fraction and you are going to have the supernatant fraction. This bead fraction is actually going to contain no messenger RNA because the messenger RNA is already been eluted.

So they are actually going to be present in the supernatant and that you can actually be able to collect into the new tube and that actually is going to give you all the messenger RNA which are actually going to have the poly A tail. So most of the messenger RNA are actually going to have the poly A tail. So now you have isolated the messenger RNA total messenger RNA from the cell and you are going to use that for synthesis of the cDNA library

or cDNA of these messenger RNA which means you are going to synthesize the complementary DNA. So in the step 2 you are going to synthesize the complementary DNA synthesis. The complementary DNA synthesis is a three step process.

In the step 1 you are going to first synthesize the reverse transcriptase. First you are going to synthesize the first strand with the help of the reverse transcriptase. Then in the step 2 you are going to remove the RNA template and then in the third step you are going to synthesize the second strand which means first you have the messenger RNA poly A tail. So you are going to use that and you are going to put reverse transcriptase reactions and so you are going to put the reverse transcriptase reaction and that is how it is going to give you the two strands. So this is going to be the RNA, this is going to be the DNA.

So it is actually going to give you the DNA which is complementary to this and then you are actually going to catalyze the reaction so that it is actually RNA is going to degrade. So you are going to do the RNA degradation by many type of reactions and that is how you are going to have the RNA. It is only the DNA strand 1 and now using this DNA strand 1 you are going to put it for the polymerization reaction and the polymerization reaction is going to give you the two strands. So this is going to be the strand 1 and this is going to be the strand 2 and this is what is called as complementary DNA and this complementary DNA then can be cloned into a suitable vector and that is how you are going to get the cDNA library. So let us first discuss about how you are going to synthesize the complementary DNA.

So we have multiple approaches one what one can use to perform the cDNA synthesis. So in the method 1 you can use the homopolymer tailing method. In a homopolymer tailing method what you have is you have this is the messenger RNA what you have and what you can do is you can just put the oligo-dt primers. So what will happen is when you put the oligo-dt primers they will come and sit next to the poly A tail. So they will sit on the poly A tail and then you can actually be able to use the reverse transcriptase and the reverse transcriptase is actually going to synthesize the first strand.

So you are going to have the reverse transcriptase and you will going to add the four nucleotides and all the four nucleotides and this is going to work as a primer and that is how you are going to have the synthesis of the first strand. So this is the first strand synthesis. Once this is done then you can actually be able to get this and then you are going to perform the alkaline sucrose gradient. When you do the alkaline sucrose gradient it is actually going to hydrolyze the RNA and it is actually going to give you the first strand the c-densin synthesis and then what you are going to do is you are going to add oligo-g column. So you are going to add the oligo-dg columns and you are also going to add the reverse transcriptase and as well as all the four dNTPs and that is how what is happen is that it is actually going to all these all the GG is actually going to bind to the CCC on which is present on the first strand and that is how it is actually going to start the synthesis of the second strand and that is how you are going to get the duplex cDNA from the messenger RNA and that you can actually be able to insert into the vector by using the suitable restriction

enzymes or you can actually be able to use this polyT and the polyC sequences as well.

So you have two choices here either you can use the linkers or you can use the restriction enzymes. So in the homopolymer tailing this method exploits the presence of polyA tail present on the messenger RNA to synthesize the first strand followed by the degradation of the messenger RNA template and the second synthesis of second strand. So it has the following step. In the step one oligonucleotide dt primer is used with messenger RNA as a template to prepare the first strand of DNA with the help of the reverse transcriptase and the dnt piece. Once the first strand synthesis is over you can actually be able to do the terminal transferase is used then the terminal transferase is used to add the c nucleotide on the 3 prime of both the messenger RNA and as well as the newly synthesized strand of the DNA.

So after this you are going to add the you are going to run the terminal transferase enzyme and that is actually going to add the ccc on both the messenger RNA and as well as the cDNA. And then this DNA RNA hybrid is loaded onto alkaline sucrose gradient. So alkaline sucrose gradient is actually going to contain the NaOH and it is going to have the sucrose gradient. So what will happen is when you are going to load this RNA DNA hybrid onto alkaline sucrose gradient. So when you load this and it actually contains the NaOH.

So NaOH is actually going to degrade the RNA because it is actually going to add on act onto the 2 prime of hydroxyl and that is how it is actually going to form a cyclized product and that is how it is actually going to degrade the RNA. Whereas DNA does not contain the 2 prime hydroxyl and that is how it is not susceptible for the alkaline lysis. So once you do the alkaline lysis and you are going to load this complex onto the sucrose gradient the RNA is going to be degraded and the DNA the first strand of the DNA is actually going to be can be isolated after the gradient. So this step will hydrolyze the RNA and it will allow the full recovery of the cDNA. Once you got the first strand of the cDNA then you actually can use oligoDg primer and you can use the cDNA as a template to prepare the second strand of DNA with the help of the reverse transcriptase and the dNTPs.

At this stage the you know the second step synthesis you have the choice you can use the reverse transcriptase and polyD polyG primers or you can actually be able to use the tag DNA polymerase you can use the poly DNA polymerase as well and you can actually be able to use that with the dG primers. So either of that can be work as and which will give you the full length cDNA double standard cDNA DNA and that can further be inserted into the vector either by the homopolymer training or by the linkers. Then you have the step 2 the method 2, method 2 is called as the Gubber Hoffman method. So Gubber Hoffman method in this approach first strand synthesis using oligoD primer in the presence of reverse transcriptase and then dNTPs then DNA RNA hybrid is treated with the RNase H to produce the nick at the multiple site. Then the DNA polymerase is used to perform the DNA synthesis using the multiple fragments of RNA as a primer to synthesize the new DNA strand.

This method produces the blunt and duplex DNA. So in this first step is same as the homopolymer telling that you are going to add the quality primers and that is actually and in the with the help of the reverse transcriptase it is actually going to give you the messenger RNA and as well as double standard RNA. So it is going to give you the messenger RNA and as well as the first strand of the DNA and now what you are going to do is you are going to add the RNase H and you are going to add the random primers and DNA polymerase. So once you add the RNase H it is actually going to chew the RNA at multiple places. This means it is actually going to add the primers at multiple places.

So it is going to leave some amount of RNA and it is going to keep some nicks. So because of that this sequence is actually going to be used. In the second step what you are going to do is you are going to use these sequences for with the help of the DNA polymerase. So when you add the DNA polymerase to this along with the dNTPs plus dNTPs what will happen is that it is actually going to use this as a primer and that is how it is actually going to start synthesis. And you know that when the DNA polymerase will run it is actually going to remove this particular sequence and it is going to synthesize its own sequence and that is how it is actually going to synthesize the new DNA strand.

So there will be no RNA present. And the same is true for this one also. And that is how you are going to get the duplex cDNA and this duplex cDNA then can further be ligated or inserted into the vector either by the with the help of the linkers or the adaptor proteins. So this is all about that how you are actually going to prepare the genomic library or the cDNA library. And once you prepare the genomic library and cDNA library you are going to get the number of clones. And then the next task is that you are actually going to do the screening of these clone with the help of the different types of analytical tools.

These tools can vary and depends on the what kind of diagnostic probe you have. So for example if you have a gene fragment right if you have a fragment of the DNA which is known that it is actually going to give you the that particular enzyme or suppose you have a antibody or suppose you have some kind of activity which is actually be associated with the unidentified enzyme X then all these things can be used for screening. So in the screening you have three options either you can use as DNA probe and you can actually be able to use that for screening the clones or you can actually have the antibodies. You have antibody which is recognizing a particular enzyme or you can actually have the enzyme activity. So either of these three methods can be used.

So if you have the enzyme activity you can actually be able to use that also to identify the clone of your interest. So let's discuss about the screening of the genomic library or the cDNA library and what are the different approaches you can use. So as I said you know you can have the two three choices either you can use the DNA sequence. So this property can be used to search both the genomic library and as well as the cDNA library to identify the gene or the clone of your interest. Then we have the approach number two the expression

of a particular protein with the immunogenic epitope site.

So this property can be partially useful to screen genomic library due to the truncation of a full gene or no expression of a gene fragment but this approach will suit to the cDNA library. So if you have antibodies which recognizes the protein of interest or the enzyme of your interest. For example if you know that in a particular pathological conditions this particular antibody is being produced in the patient. So you want to identify if you want to identify the enzyme what you can do is you can take this antibody and you can prepare a genomic library or the cDNA library and then you can use that as a probe to identify the clone. In this approach the cDNA library is more suitable because cDNA library means you are actually going to have the expression clones.

So in the expression clones the clones are actually going to start suppose this is a clone and this is the gene what you have inserted this is the messenger RNA the cDNA it is actually going to produce the protein and this protein is then can be detected with the help of the antibody. Whereas in the case of genomic library the problem is that genomic library sometime may have the truncated proteins. So sometime it may have half protein and the half gene may be of the other. So in that case it may actually give you the protein and that time you can be able to use but if it does not if it only gives you a truncated protein for example if it only gives you a half protein and that half protein does not contain the antigenic site then in that case it is actually not going to work. So 100% if you have the antibody which is going to use as a tool to recognize the enzyme you can be able to use the cDNA library.

But for the other case like the genomic library the DNA is more suitable. Then we have the enzymatic activity for example if you are trying to explore enzyme which is associated with a particular activity but you do not know the gene in that approach you can actually be able to use the enzymatic activity. So this property exploits the ability of a protein fragment to exhibit enzymatic activity it is useful for the screening of cDNA library but it is not useful for the genomic library because of the simple reason that the genomic library may or may not be complete. So the gene fragments are not complete they may give you the truncated proteins and those truncated proteins may or may not give you the activity.

So let us start the first method. The first method is where you are going to use the DNA sequence or DNA probe for screening the clones and that can be used both for genomic library and as well as for the cDNA library. So DNA if you see the structure of a DNA, DNA has the double helical structure where you have the nucleotides what is present inside the helix and these nucleotides have the very peculiar base pairing. You always know that the adenine is always making a pair with thymine and whereas the guanine is always making a pair with cytosine. So because they are very strict and they are also only making a pair you can actually be able to use that as a sequence. So wherever you have the A you are going to have the T on the template wherever you have the G on the probe it is actually going to have the C on the template.

So suppose I have a template DNA or if I have the genomic sequence which I want to screen then what I will do is I will prepare a probe like for example I have prepared a probe like this. So this is the probe I have prepared. Because this is the sequence I know that it is actually going to bind to that particular gene which is responsible for the production of this particular gene. So now when it is actually going to recognize it is actually going to recognize a protein or the DNA. What will what DNA sequence it is going to identify? It is actually going to identify a DNA sequence of this.

So wherever it will find a DNA sequence with this it is actually go and bind and that is how you can be able to identify this template DNA or this cDNA clone or the clone DNA with the help of the probe. So this is going to be the probe this is going to be this and the probe will where the probe is binding for that you have to put some kind of you know the tag actually. So you can actually be able to put the fluorescent tag or you can actually be able to put the radioactive probe or radioactive tag. So if you add the radioactive probe which has this sequence it will go and bind to all the DNA sequences or the gene sequences where you have the this particular sequence present. Now how you are going to do this? You are going to use this with the help of the from cDNA library.

So this is suppose this is the cDNA this is the library whether it is the genomic library or cDNA library. So you are going to have the main plate or the plate where you are going to have all the clones and that plate is called as the master plate which means this is the original plate where the your clones are present. So imagine that these are the clones you have. So then first in the step one what you are going to do is you are going to first transfer the master plate and you are going to prepare a replica plate. So that you can actually be able to work with all the clones without destroying the master plate without destroying the original clones.

So you are first going to prepare a replica plate. So you are going to just insert invert this onto another plate and that is how it is actually going to give you the replica plate or you can actually be able to transfer that onto a nitrocellulose membrane and that is how it is actually going to give you a impression of the clones onto the nitrocellulose membrane. Now in the step three the step two you are going to do the lysis of these cells. So once you lysed the cells it is actually going to denature the DNA and it will actually going to bind the matrix which means all the clones are all these are you know cells. So they will be get lysed and that is how the DNA will come out. So DNA will come out from these cells but they will not going to washed away because they will go and bind to this nitrocellulose membrane.

Now what you are going to do is in the step three you are going to add the DNA probe which actually has the tag. So either it can have the radioactive tag or is a fluorescent tag. So once the tag is there it will actually go and bind to its specific you know for example if it is binding to this particular clone and then what you can do is you can just take this replica plate or you can take this replica membrane and then you can actually compare that with

the help of the master plate and you will know that this is the clone what is response or which is the where the gene of my interest is present. And then what you can do is you can just take out this gene of interest and you can just grow them into the media and that is how you can be able to isolate the plasmid or you can actually isolate the recombinant DNA and from this plasmid you can be able to isolate the gene of your interest and that is how you can use this gene for further downstream applications.

So it has the following steps okay. The step one you are going to prepare a suitable radioactive probe you can prepare also the fluorescent probe then you are going to prepare a replica plate. So this is what the replica plate then you are going to transfer of the colonies on the nitrocellulose membrane then you hybridize that with a specific probe which means the radioactive probe and then you are going to wash and development of the membrane by the audio radiography and that is actually going to tell you on which colony it is binding and that is how you can actually go back to the master plate and you can actually identify. Now how you are going to prepare the radioactive probes so you can actually be able to use the multiple methods of preparing the radioactive probes. So in the preparation of the radioactive probes you can use the random primer methods. So in the random primer methods in this method random primer is used to anneal to the template and then a PCR reaction is performed in the presence of the radiolabeled nucleotide after the PCR the newly synthesized DNA strand is labeled with the radioactive nucleotides.

So what you are going to do is suppose this is the template for which you want to synthesize the probe so what you are going to do is you are going to add oligonucleotide primers and you will do the hybridization so it will actually go and hybridize and then what you are going to do is you are going to do a DNA synthesis with the help of the clinofragments and the 4 dNTPs so it will actually going to synthesize the strands and that is how it is actually going to incorporate the radioactive nucleotides and that is how it is actually going to produce the labeled probe and that labeled probe you can actually be able to purify with the help of the gel filtration chromatography and that can be used for further downstream applications. Then we also have the terminal transferase method so we also have the terminal transferase method so in the terminal transferase method you are going to use an enzyme which is called as terminal transferase so in this method a terminal transferase enzyme will label the probe at the end of the last nucleotide of the probe. Probe is incubated with the labeled nucleotide and the terminal transferase enzyme and will add the labeled nucleotide at the end. A partial purification of the gel filtration column will give you the labeled probe. So for example this is the gene so what you are going to do is first you are going to rotate it with the alpha endonuclease and that is actually going to cause the nicks and then you are going to have the terminal transferase and as well as the radioactive DATPs for example if I want to you know radio labeled DATP so what will happen is it is actually going to add the A's on one end and that is how it is actually going to incorporate the radio activity in both the strands of this particular gene and that is how I am going to get the radioactive probe with one end of the probe as radioactive and that I can further purify with the help of the gel filtration protocol.

And now once the probe is ready I can use that for the screening purpose. So for the screening purpose I have to prepare the replica plate so as the original genomic DNA or cDNA library is precious and will be consuming in the later stage all procedure is performed with the replica plate containing the colonies in a identical manner. First you are going to transfer so the clone is transferred on to a nitrocellulose membrane with retaining the identical pattern of the colonies on the master plate. The cells on the membranes are lysed and the released DNA is denatured, deprotonated and allowed to bind to the membrane. So this step is very crucial and it actually going to decide what will be the what will be the success of your screening because if you could be able to do this successfully then you what you are going to do is you are going to actually lyse the cells and the DNA is actually going to immobilized to the site of that particular clone.

Then you are going to do the hybridization so available probe prepared in step one will actually going to be added the probe will add to the target DNA due to the base pairing the membrane is washed to remove the unbound and then you are going to do the development of a durogram. So this position of the durogram is detected by the autoradogram the position of the signal on the membrane can be washed with the master plate to get the location. For example if this is the plate and this is my master plate so I will actually super impose both of them and then I will know that okay this is the clone which actually is giving the signal or the durogram. So then I can actually be able to isolate this and I can isolate the gene of interest. Now the second step second method so second method is the screening by the immunological method.

So in the immunological method you are going to use the antibody as the probe. So antibodies can be tagged with the enzyme or you can actually be able to add the fluorescent. So antibody can be tagged to the fluorescent dye or it can be attacked to the enzyme such as HRP or alkaline phosphatase. So in this case what you are going to do is step one is same you are actually going to prepare first master plate and then you are going to prepare the replica plate and from the replica plate you are going to transfer the cells and onto the membrane and that is how you are going to prepare the netocellulose membrane and then you are going to light the cells and allow the protein to bind to their site. And once that is done then you are going to add the primary antibodies and once you add the primary antibody the primary antibody is actually will go and bind to the sites wherever you have the antigen of interest.

You treat the matrix with the primary antibody and the primary antibody is actually going to you know bind the proteins what is present within the cell and that is how it is actually going to give you and then what you are going to do is you are going to add the secondary antibodies. So secondary antibody is actually going to be tagged with the enzyme or the fluorescence right so that secondary antibody will go and bind to the primary antibody and wherever it will bind it is actually going to give you a signal. So looking at the signal right for example this is the clone which is giving you the signal you can go back to the original

plate so original plate is saying that this is the clone from where I am getting the signal. So in that case you actually can use that and you can isolate that clone and you can actually be able to grow that into the large quantities and that is how you can actually be able to isolate the plasmids and from the plasmid you can be able to isolate the gene and that gene you can actually be able to use for the over expression purpose. So these are the two methods which are very very popular either the immunological method where you are going to use the antibody as a probe or the DNA probes.

Apart from that you can also use the screening by the enzymatic method. So this method is based on the ability of a protein to exhibit an enzymatic activity. This method is not very specific but it allows us to identify a class of protein with the known enzymatic activity. So in this case also you are going to have the same steps what you are going to do is you are going to first prepare the master plate right.

So this is the master plate. So this is the master plate from the master plate you are going to prepare the replica plate like this okay and this is going to be your replica plate and replica plate from replica plate you are going to first transfer that onto a nitrocellulose membrane. So you are going to prepare that onto the NC membrane and on this NC membrane you are going to add the substrate for your enzyme. So you are going to lyse the cells right you are going to lyse the cells and you are going to add the substrate. So once you add the substrate it is actually going to give you the signal okay. This signal you can actually be able to compare which is there on the replica plate to the master plate and that is why you will say that okay this is the clone of my interest and that is why it is actually going to give you.

The only issue with the screening by the enzymatic method is that it may not be unique because in some cases you might have to see that multiple clones are actually going to give you the activity because the substrate is a very very non-specific probe because substrate can be used by the multiple enzymes and that is how it may actually give misguide you in terms of the getting the clones. For example if I use the glucose as a substrate right so glucose can be used by the hexokinase glucose can be used by the glucokinase glucose can be used in any other reactions also. So in that case you may get the clone of your interest but it is not very specific and you may actually be able to use or you might have to use the other screening method to further verify the clones. So this is all about how you can be able to screen the genomic library and as well as the cDNA library and once you actually have screened and you say that okay this is the clone I have to isolate you can actually have to perform the multiple steps to isolate this particular clone also. So how you are going to isolate the gene so once the position of a clone is known it is extracted from the master plate and the plasmid is isolated in few cases the clone is further diluted to check the homogeneity of the clone the purity of the clone and the presence of clone is further tested with the PCR using the size sequence specific primers.

So what happens is that you what you are going to do is you are going to first isolate the

clone okay and then you are actually going to dilute that into 1 is to 10 or 1 is to 100 dilutions okay and then again you are going to plate that onto a plate okay so in that case you are going to get all the colonies okay. Ideally if this clone is pure which means it only has a single gene it actually will all these clones are identical to each other. So again if you repeat the probing reactions like either you use the DNA probe or antibody probe or the enzymatic method all these clones should actually give you the signal which means all these clones should give you the signal because if it could also happen that some of the clones actually will give you the signal okay. In those cases what happen is in that case the this particular clone is not pure it may have actually the multiple clones which are coming together or when you are isolating the clone you actually got the cross contamination from the neighboring clones also. So in that case you have to first do the you know this is the primary screening through which you got this clone then you might have to do the secondary screening by further diluting these clones and doing the same reaction again and again and you have to repeat that until you are actually getting all the clones which are actually going to give you the signal.

So that is what you have to do if you want to isolate the gene. So once you got this you can actually be able to isolate the plasmid and that plasmid is actually going to contain the gene of your interest so that you can actually be able to use the gene and then you can actually be able to use the PCR and you can actually we get the amplified gene and that amplified gene you can put into the expression clones or expression vector and that is why you can be able to use that for protein production. So this is what we have discussed. So we have discussed about the approach one where which is very very common or very very much popular into the pre genomic era when the genomic sequence were not known and you were only knowing that there is an antibody which has been found in the patient or there is a genomic sequence or the DNA fragments is found and something like that. So in those cases you have to use the genomic library so you have to take the genome you have to prepare the genomic library or you might have to isolate the messenger RNA and you have to prepare the cDNA library and then once you are you prepare the genomic library or the cDNA library you can use that for screening the clone of your interest with the help of either the DNA probe or the antibodies or the enzymatic method. Once utilizing these three screening tools you can be able to screen the clone of your interest and once you got the clone of interest you can just isolate the plasmid you can prepare the you can isolate the gene of your interest you can use the PCR to amplify the gene of your interest and then you can actually clone it into the suitable expression vector and that is how you are going to get that particular gene responsible for the production of the enzyme into a expression vector and that is how you can be able to use that for enzyme production.

So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss about the second approach where the gene sequence is known and that is very common and very popular nowadays for the enzyme production. So with this I would like to conclude my lecture here. Thank you. you

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 10

Molecular Cloning

Lecture-42 Cloning Vectors

DNA selection marker and the bacterial origin of the applications before vector delivery to yeast it is digested with the unique restriction endonuclease to produce linear DNA to increase the transformation efficiency and the integrations.

In most of the cases integration is done in such a way that the yeast chromosome DNA remain intact and integration may not affect the yeast growth. But in an alternate approach a portion of the yeast chromosome DNA is replaced with the vector DNA through the homologous recombination. These vectors are known as the transplant integration vector and they have the foreign DNA selection marker and homologous DNA into the region of chromosomal DNA to be replaced. Then we can also have the third type of vector which is called as a yeast artificial chromosome or YAC vector.

So, yeast artificial chromosome is the vector of the choice used for cloning very large DNA fragments. Remember that when we were talking about the preparation of the human genomic library we have said that we are going to clone that into the YAC vectors. To prepare the genomic library YAC vector is like a chromosomal and it has the ARS sequences centromere sequences and telomeres at the two end to give the stability. It has an MPC resistance genes for selection in E. coli and an E.

coli in the E. coli in the original replication for the propagation in bacteria. In addition it has the ARS for replication, SEND for the centromere functions and the URA3 and TIP211 for the selection in the yeast. For cloning YAC is digested with the enzyme called SM1 and BAMH1.

So, you are going to digest the YAC with the SM and BAMH1 and you are going to treat that with alkaline phosphatase to generate a linear plasmid. Now foreign DNA is added for the ligation. So, at this stage you are going to add the foreign DNA and that's how it is actually going to get inserted into these two fragments and the recombinant DNA will allow a yeast to grow on a uracil and tryptophan deficient media. So, screening anyway we are going to discuss in our subsequent class. Then we have the eukaryotic plasmid, another eukaryotic plasmid which you can use in the insect cell lines.

So, these are the vector for expressing the protein in the insect cell lines and or the baculovectors right. So, baculovirus is a rod-shaped virus infecting the invertible bait including the insect cells. Post infection the virus is either released as a free viron or many virus particles are trapped in a protein complex known as polyhedron. The protein responsible for this is polyhedron. The protein responsible for trapping the virus into polyhedron is polyhedron and it helps in the transmission of virus from one host to another.

The polyhedron is not important for virus propagation, but it is under very strong promoter to produce the protein in large quantity. So, realizing this fact the replacement of the polyhedron gene with a foreign DNA fragment will allow the expression of the protein in a large quantities. So, the baculovirus autografa califonia multiple nuclear polyhedron virus or the ACMPNP this is what is used as a vector to express the protein. The transfer vector map of the ACMPNP is given right. This is the vector the map of the ACMNPB where you have the cloning site.

This is the cloning site what you have the gene of interest will be inserted into the cloning site place adjacent to the promoter. So, you have the this is the polyhedron gene promoter right and it has the polyhedron termination sequences downstream to the cloning site to stop the transcription of the clone gene and or more will be discussed in future lectures. So, this is what you have you have the ACMPNP vector where you have the polyhedron promoters and next to the promoter you have the cloning site. So, within this cloning site you can actually be able to insert the gene of your interest right and that's how it is going to start expressing this particular protein instead of the polyhedron polyhedron and it also has the termination sequences. So, that the transcription is going to stop after this and you can actually be able to take this and put it into the insect cell lines and that's how you're going to express.

Then we have the eukaryotic vectors like the mammalian vector. So, large number of excellent mammalian vectors are in circulation to clone the eukaryotic gene for the protein synthesis and to study the transcription mechanisms. It contains a eukaryotic replication origin from an animal virus such as SV40 from a simian virus a promoter to derive the expression of foreign gene and the selection marker and the other eukaryotic features such as adenylation, transcription termination, etc. So, this is what is the mammalian expression vector where you have the multiple cloning site, you have the promoter and you have the other features of the plasmids. You can also have the origin of replication for the eukaryotic system.

You can also have the origin of replication for the bacterial system. Then you have the antibiotic resistance genes and so on. Then we have the bacteriophage based vectors. So, bacteriophage lambda based vectors are the virus using the bacteria as their host for replication. Bacteriophage lambda is the virus of E.

coli and have been used to develop vector for the genetic recombination. So, what you have is a bacteriophage genome. So, far genome is a linear double standard DNA of 48.5 kB. On both end of the genome, it has a stretch of 12 nucleotides which are complementary to each other.

So, you have the two sites on both end of the genome. One is called as the left cohesive site, the other one is called as the right cohesive site and within this you have the different region of the genome which is expressing for the different part of the body. So, these sites are called as cosites and it allows the circulation of the viral genome after entering into the host cells. Genes are arranged between these two cohesive sites of the code for the protein responsible for making the head, tail, factor for recombination and the process of lysosin. The central region of the genome is non-essential and can be replaced without much affecting the growth and the infectivity of the virus.

As a result, this region can be exploited to develop a cloning vector with multiple approaches. So, how it is actually going to pack the genome? So, the far genome is replicated by a rolling circuit model to produce a long genome whereas cosites are present on the regular interval, right? So, when it is actually going to produce the far genome into the outside, right? And it is actually going to start producing the cosites. So, you can imagine that you have one cosite, you can have another cosite and once these two cosites are going to be out, they are actually going to come together and they will get circle. The two flanking cosites and the DNA between them constitute the viral genome or the monomeric unit. In the presence of head precursor, the long genome is cleaved into the monomeric unit and encapsulated.

Nicks are introduced on both the strand of the genome to generate the linear strand to serve as a cohesive site to facilitate the circularization in the host. Okay, so this is what. The bacteriophage lambda cloning vector has a middle segment responsible for the insertion or excision and this region can be replaced with a foreign DNA with the help of the two BMH1 sites present on the either side of the insertion or excision regions. Lysosin cycle, lytic cycle and it will form the plaques. So, bacteriophage vector are the EMBL-3 and EMBL-4.

So, this is what exactly going to happen, how you are going to insert the foreign DNA into a IE site. Okay, so this is what all about the different types of vectors what are available for the cloning of the foreign DNA or cloning of the enzyme into the suitable expression vector and depending on the production, depending on the origin of that particular gene, you can be able to have the flexibility to choose the different types of vectors, right. You can use the mammalian expression vector, you can use the *E. coli* expression vectors and so on. Now, once you have chosen the vector, so that is one thing.

The second is you have only know the enzyme, then you are actually going to use them together to generate the recombinant DNA. Now, how you are going to do that is you are

going to run the parallel multiple reactions. So, what you are going to do is first from the genome, you are actually going to generate or you are going to produce the fragment. So, you are going to produce the fragment, right. Similarly, from the vector you are going to have the circular vector.

Okay. This circular vector will have the multiple cloning site and within the multiple cloning site, you are going to have the, you are going to introduce your insert. Now, for this fragment you are going to generate or you are going to produce the restriction enzyme. So, for example, you can treat it with the two restriction enzyme, restriction 1, 2. So, it is actually going to produce the you know the fragments with the cohesive ends and these two cohesive ends like for example, this is for the restriction enzyme 2, this is for restriction enzyme 2. Same is true for the vector also.

So, vector also you are going to digest with the help of this and that, right. And, okay, so these are the restriction enzymes. So, again the same set of restriction enzyme you are going to use for cutting the vector also. Then you both took for the ligation reaction.

Okay. Just now as we have discussed in the previous lecture, right, and once you put the ligation reaction, you are going to have the recombinant DNA with your insert into the multiple cloning site. Okay. So, this is going to be your chimeric DNA or recombinant DNA. Now, this chimeric DNA you are going to transform into the host.

Okay. And the host is going to have this, right. So, since this restriction site this plasmid is going to have the origin of replication, it is actually going to replicate. So, depending on the genome, depending on the origin organisms from which you are isolating this gene, you can be able to choose the different types of vector. And depending on what kind of applications, if you want the protein or the enzyme in milligram range, you might be good to use the equalized system. If you want to use the higher level, then you can use the yeast expression vectors.

Similarly, if the genome is or the organism is very close to mammalian system, then you should use the mammalian expression system or the yeast expression system because that will give you the properly folded proteins. So, this is all about this, you know, the cloning of the particular enzyme into the suitable vectors. And now you have, what you have done is you have generated a recombinant DNA, right. So, once you have, so this is the recombinant DNA what you have generated, right. Now, once you generated the recombinant DNA, it has to be transformed into the host and then you are going to use that for the protein production or the enzyme production.

So, in our subsequent lecture, we are going to first discuss about how you are going to insert or how you are going to, you know, devise the different types of strategies to deliver the DNA, recombinant DNA into the host of your choice and subsequent to that, we are also

going to discuss about the protein production. So, with this, I would like to conclude my lecture here. Thank you. .

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 10

Molecular Cloning

Lecture-43 DNA Delivery (Part 1)

Hello everyone, this is Dr. Vishal Tewedi from Department of Biosciences and Biogeny IIT Guwahati. In the previous module, we have discussed about how you can be able to isolate a gene from the genome, whether you are aware of the genomic sequences or whether you are unaware of the genomic sequences, you can be able to use multiple approaches to isolate the gene of your interest, which is going to code for the enzyme of your interest. And then once you have got the fragment, we have also discussed in the previous lecture, how you can be able to clone that fragment into the vector of your choice. Now, once you have cloned the fragment of your choice into the vector, the vector has to be delivered into the suitable host. So, in the today's lecture, we are going to discuss about the various strategies what you can actually be able to use for delivering the DNA into the suitable host.

So, what you can see is this is a scheme for enzyme production, right. And in this scheme, we have already discussed many aspects of this scheme where we have said that the you are going to isolate the gene fragment from the genome, right, either you are aware of the genomic sequences or you are unaware of the genomic sequences. This means either you will use the genomic library approach or the cDNA approach, or you are going to use the PCR. Once you have these genomic sequences, then you are going to perform the decision digestion of this particular sequence.

And that's how you're going to get the sticky ends. Same you are going to do for the vectors, right. In this case, we have taken an example of the plasmid, which is a bacterial vector. And then you're going to put the decision digestion and you're going to have the sticky ends of the plasmid, then you're going to put them for the ligation reaction and that's how you're going to get the recombinant clone. This recombinant clone is has to be delivered into the suitable host so that you can be able to use that for protein production.

So, when we talk about the DNA delivery in host, we have to, you know, we have to understand the many aspects of this particular process. So, DNA delivery in host, so you can imagine that how the people have discovered or how the people have get to know about the DNA delivery in host. So, you can imagine that you have a donor cell from which you are actually going to have the, you know, the DNA molecules. And these DNA molecules are, you

know, that are negatively charged. So, these DNA, the charge on the DNA is negative charge on the DNA molecule interacts with the host cell, especially with the cell surface, which means if I change the cell surface chemistry of this donor cell in such a way that it is actually going to take up these DNA and they will go and you know, stick to this particular, you know, cell wall or the plasma membrane, they will be going to taken up by the cell.

So, the surface chemistry of the host cells, host cell surface chemistry either will attract or the ripple DNA as a result of the opposite or the similar charges, which means the cell surface of the host is either going to be positively charged or the negatively charged. If it is positively charged, it is actually going to attract the DNA, if it is negatively charged, it is going to ripple the DNA. So, presence of the cell wall in case of bacteria or fungus or the plant causes the additional physical barrier to the cellular uptake and the cellular entry. So, actually the entry of the DNA is not a facilitated process, it has, it has multiple barriers and multiple hurdles. And all these hurdles has to be overcome by preparing our competent cells.

So, what is mean by the competent cell is the cell which is actually going to be readily be taken up the, so which is readily going to take up the recombinant DNA. And you are supposed to prepare their competent cell, which means you are not, wild type cells are not competent enough, they are not going to take up the DNA. But you can actually be able to, you know, treat them in such a way that it is going to change the surface chemistry of the cell surface and that is how it is actually going to take up the cells. Naturally, the, this process is called as the transformation, which means the process in which the DNA fragment from the one cell is going to be taken up by the other cell. So, you can imagine that if the donor cell is resistance for the antibiotic, for example, you have taken an antibiotic and that is been, you know, some, some bacteria cells are actually acquired the resistance for that antibiotic cells by many means because of some presence of some DNA.

So, if you have a presence of DNA, which is responsible for providing the antibiotic resistance, this particular donor cell is actually going to be overcome or it is actually going to be overcome from the antibiotic mediated killing. So, in that case, what it will do is it will actually going to throw this particular DNA and this DNA is going to be very, very valuable for the other bacterial members. So, what they will do is they will going to accept by the other cells and as soon as the other cell is going to take up this DNA, it is actually going to also acquire this cell and this is actually the way in which the transformation is actually going to be you know, going to help the single bacteria to acquire the resistance and also to spread the resistance throughout the colony. How it happens? It happens that you are actually going to have the donor cell from the donor cell, the donor DNA or the DNA fragment, which is actually going to have the antibiotic resistance genes or any kind of phenotypic gene is actually going to put into the extracellular media and then this extracellular media is actually going to interact with the acceptor cells. So, you are going to have the competent recipient cells.

So, that donor DNA fragment will bind to the competent recipient cells and that is how it is actually going to be taken up. So, mechanism of transformation is the process by which the cell free DNA is taken up by the another bacteria. The DNA from the donor bacteria binds to the competent recipient cells and DNA enters into the cell. The DNA enters into the recipient cell through an uncharacterized mechanism. The DNA is integrated into the chromosomal DNA through homologous recombinations.

Naturally, transformation is common between the closely related species, which means there are you know, through natural transformation, we have taken an example where you can actually be able to use that for transferring the resistance from the resistance or the phenotype from the one cells, one bacterial population to another bacterial populations. But the in the lab, what you are going to do is you are going to add the extracellular DNA or you are going to add the recombinant clone recombinant DNA and in presence when you add this recombinant DNA to the competent recipient cells, the DNA is actually going to interact with the competent recipient cells and that is how it is actually going to taken up by the cell. And once it is going to taken up by the cell, it is actually going to go for cell division and that is how you are going to have the transform cells. So, where the copies all on the both side of the DNA. Some time you actually going to have the integration sites on these DNA.

So, they will integrate into the genome, otherwise they will remain as the extra chromosomal DNA in the form of the plasmids. Now, how you are going to prepare the competent cells for your laboratory experiments. So, you can use the depending on the bacterial species, you can be able to use the different types of chemical reagents which are actually going to change the surface chemistry and will make the cell competent so that they will take up the new cells, they will take up the the extra cellular DNA. So, for example, in the case of tetra focus new in new many, you are going to you are going to treat the cells with the mitomycin C or fluoroquinol. So, when you treat the cells with mitomycin C or fluoroquinol, they are going to be competent and they will actually going to take up the extra cellular DNA.

Similarly, you have the basillus subtilis, you are going to treat the cells with the UV light and that actually is going to induce or that is going to change the surface chemistry and that is how it is actually going to be competent cells. In the case of helicobacter pylori, the bacteria which is responsible for the ulcer or is going to be treat with the ciprofloxacin, ciprofloxacin is an antibiotic. So, the treatments are different because your surface chemistry is going to be different, the physiology of that particular bacteria is going to be different. So, that is why you are actually going to treat it with the different competent agents. Then we have the dibunopila pneumophila, then that is actually going to be make the competent by treating you with the mitomycin C, norfoxacin, ophloxacin and all these kind of antibiotics and hydroxyurea and as well as UV light.

For example, E. coli which is called as laboratory strain. So, this is the lab strain and that

can be competent by treating with the covalent, you know, the covalent, covalent, you know, chemicals like calcium chloride and rubidium chloride and the cells are how you are going to make the competent cells. So, you are going to treat the cells with the different agents, so all of these agents. So, what we are going, so just for sake of how we can be able to make the competent cells, I am giving you an example of the E.

coli cells. So, we in the E. coli cells, we are going to take an example of how you can be able to use the calcium chloride to prepare the E. coli competent cells. So, preparation of the competent cells, so bacteria is incubated with the divalent cation like the calcium chloride, manganese chloride or the rubidium chloride for 30 minutes at 4 degrees Celsius. And what, so bacteria you are going to, first you are going to grow the bacteria and then your bacteria is going to incubate it with a divalent cation such as calcium chloride, manganese chloride or rubidium chloride for 30 minutes at 4 degrees Celsius.

What will happen in this period? During this process, the cell wall of the treated bacteria is going to swell and it gathers the factor required for the intake of the DNA docked onto the plasma membrane. So, you can imagine that when you are doing all this preparation, it is actually going to cell swell, it is actually going to take up the calcium chloride or the divalent cations. And these divalent cations, what they are going to do is they are actually going to take, they are going to make the cells a little fragile. So, in that case, you might have to take care of these cells very nicely because if you, they also get susceptible for any kind of shear stress. So, shear stress is a, it means that you are actually going to, suppose when you are swimming into the water, right, the water, whatever the stress you will feel when the water is actually hitting you, that is called as a shear stress.

Shear stress means you have two layers and these two layers are actually rubbing to each other and that is actually being responsible for causing a friction and that is actually called as shear stress. So, they are very susceptible for shear stress. For example, if you spin them a very high speed, you probably will actually go into live cells. So, that is why these computational cells are very fragile, you cannot run that very high speeds, okay, and you also cannot do the pipetting at a very high pipetting. Like for example, if you do a pipetting a lot of, you know, with the thin bore tips, it is actually going to destroy the cell because it is going to live the cells.

How you are going to store the competent cells? You can, once your competent cell preparation is over, then you can actually be able to add 15 to 20% glycerol and you can store it at minus 80 degrees Celsius and whenever you require, so 15 to 20% glycerol is actually going to work as anti-freezing agent, so it is not going to allow the formation of the water crystal and that is how it is actually going to protect the bacteria from getting any kind of damages. So, you have to be very careful that the cells are very fragile, they are sensitive for the centrifugations and they are also sensitive for the pipetting. That is why you can, once you prepare the competent cells, you aliquot them in into a suitable volume and then you can just use that same, that volume completely. You cannot just pipette it out

and prepare the more aliquots. How you are going to prepare the competent cells? So, what you can do is, first you are going to do, grow the bacteria, okay.

So, if you see the bacterial growth curve, what you are going to see is, it has the lag phase, it is going to have lock phase, it has the stationary phase and it has the death phase or the decline phase. So, the cells which are in the lock phase, stationary phase or the cells which are in the decline phase are actually under the extreme stress, which means these three phases of the cell, right, whether you are under the lock phase or the stationary phase or the death phase are not suitable for preparing the competent cells. So, you have to grow the bacteria in such a way that it should be in a lock phase, in a logarithmic phase, so that they are very healthy, they are actually acquiring the different types of factors and there is no dearth of nutrition. So, there is no loss of nutrition. So, they are actually going to have a lot of nutrition and they are actually having the ability to produce a lot of factors.

So, that is actually going to be the best bacterial culture what you are going to use. So, the growth stage of the bacteria has a significant impact for its ability to take up the foreign DNA. The bacteria at lock phase is more active and efficient to perform the DNA damage and repair than the stationary phase. As a result, it is preferred to use a bacteria of lock phase for making the competent cells for the transformation. So, once you grow the bacteria, bring it to the lock phase and then you, you know, collect the lock phase bacteria and then you are going to prepare that using, use that for the competent cell preparation.

So, how you are going to do the transformation? So, on the day of transformation, competent cells are incubated with the DNA or the circular plasmid containing appropriate resistance gene, such as MPC in resistant genes for the 30 minutes on ice. So, first you know, what you are going to do is, this is the step one number one. So, in the step one, what you are going to do is you are going to thaw the competent cells, right? Because remember that the competent cells are going to be stored at minus 80 degrees Celsius in the 15 to 20% glycerol, right? So, you are going to first what, what you are going to do is you are going to thaw the cells and you are going to use that for transformations. Then in the step one, you are going to incubate the DNA, either the linear DNA or the circular plasmid within the ice, right? Along with that for 30 minutes. So, and then you are going to add the pestle in the resistance, okay? And then once this step is over, so you are going to have the 30 minutes incubation and once the 30 minutes incubation is over, then you are going to give the heat shock.

So, in the step two, you are going to do a heat shock. So, competent cells are given a brief heat shock such as 42 degrees Celsius for 90 seconds to relax the cell wall and high temperatures thus causes upregulation of the factor responsible for DNA recombination and repair. So, then what you are going to do is you are going to give the heat shock. So, in the heat shock step, what will happen is that because you are giving the heat shock, the cell is actually going to swell and it is actually going to take up the DNA into that. And because you are providing the heat shock, it also going to induce the production of the cellular

machinery which is a sample for DNA recombination and repair because that is very, very important event.

Because if the even if it has taken up the DNA and it cannot actually be able to recover from the damages, then the cell is actually going to die. Then in the next step, what you are going to do is you are going to add once the heat shock is over, then you are going to add a chilled media and that is actually going to you know that is going to bring the rigidity of the cell wall or the plasma membrane. And that is how it is actually going to seal the pores, what is going to be prepared into the plasma membrane and that is how it is actually going to help in terms of the faster recovery of the transform cells. So, once you add the chilled media, you can actually allow them to recover for 30 to 45 minutes at 37 degrees Celsius. And after this recovery, you are actually going to plate the cells onto the agarose plates.

So, after this, it is plated onto the solid media with the appropriate antibiotics such as empicillin and allowed to grow for another 18 to 24 degrees Celsius at 37 degrees Celsius incubator. So, when you are done with the chilled media, you added the chilled media, you let them to remain into the 37 for some time and then you are actually going to plate them onto the agarose Lb empicillin plate. So, for example, in this case, we have taken the empicillin resistance gene. So, when you do the empicillin plating, which is actually and then you incubate into the incubator for 18 to 24 hours, then it is actually going to form the colonies. And now these colonies are the transform colonies which are actually going to give you the which has taken up the DNA.

Because the untransformed colonies will actually not going to have the empicillin resistance and that is what they will actually going to die in this Lb empicillin plate. So, transform cells with appropriate resistance will grow and it will give you the colonies. So, once you are done with the transformations, how the bacterial plates will look like. So, this is actually the untransformed or the control plate.

And this is your transform plates. So, what you see here is, we have transformed a plasmid into the Lb empicillin. And if you count the number of colonies, it is 80 to 90 colonies what we got into this particular plate, whereas there is no colony in the control plate, which means the plate where we have not done any plating, we only contains the cells it does not contain that. So, this is without DNA and this actually has plus DNA. Now, how using this in use this number, how many colonies you got, you can be able to calculate the transformation efficiency. What is mean by transformation efficiency? The number of colony forming unit obtained by transforming 1 microgram of DNA into a given volume of competent cells.

So, number of colonies what you got from the 1 microgram of DNA is called as the transformation efficiency and it is a very, very important parameter to judge how good your competent cells are because if they are very bad, you cannot use them for the cloning reactions. For example, if you transformed the 1 microliter of 0.09 microgram per

microliter plasmid into 100 microliter of competent cells, you added 900 microliter of LB to your cell to get a total reaction volume of 1000 microliter and then plated 100 microliter of the transformation. The plate has 450 colony on its next day, which means you have actually added this much amount of DNA and this much amount was the reaction volume. So, what you are going to do is, you are first going to calculate the amount of DNA what is you have used.

So, if you calculate this, what you are going to know is that you have plated 0.001 nanogram of DNA. Now, if you want to calculate the efficiency 450 colonies from 0.005 nanograms into 1000 because this is what the dilution you have done and that will actually going to tell you that you have a transformation efficiency that is 4.

5 into 10 to the power 8. So, that is very good actually. So, 10 to the power 8 is a very good efficiency. Sometime you may get even more 10 to the power 11 and so on. So, that is going to be very, very good actually. So, any number which is above 10 to the power 4 or 5 is reasonably okay.

Apart from the transformation by the chemical method, you can also use you can also do the electroporation. So, electroporation is a method where you are going to use the electrical pulses into the cell. So, what will happen is that when you put the electrical pulses, you are actually going to make the hole into the plasma membrane and utilize the and since the DNA is already out, it is actually going to enter into the cell and that is all and after that you are going to put the chilled media and it is actually going to resell this pore and that is how it is actually going to be taken up the DNA. So, the plasma membrane is composed of the lipid and protein. These macromolecules give a partial conductance to the cell membrane.

So, when a high electrical pulse is given to the cell, the charge runs across the membrane and partially disturb the arrangement of the lipid molecule. As a result, it makes the formation of pore and allow easy passage of the macromolecule, especially the charged molecule like DNA into the cell. After the electroporation, the cell is allowed to recover from the damage and it forms a colony on the selective solid media. So, you can see that this is a plasma membrane.

So, plasma membrane is made up of the lipids. So, lipids are arranged like this and lipids are partially being charged. So, when the plasma when there is a very high pulse which goes on top of this, these lipid membranes are lipid molecules are getting rearranged and because of that, it actually allows the passage of the DNA into the cell. And this DNA is so this is actually very momentarily. So, once you add the media, this pulse is going to be over and that is how it is actually going to seal this particular thing. And that is how the DNA is going to be delivered into the host.

And then later on, you can actually allow the cells to recover and that is how you are the

transformation done. The advantage of electroporation is that it does not depend on the surface chemistry of the cell. It depends on the so that is why it is very robust compared to the chemical transformations. The disadvantage is that the electroporation is required a specialized instrument like the electroporator and it also requires the electroporeses. So, what are the factors affecting the transformation efficiency? The plasmid size, then the form of DNA, you can know that when we were discussing about the plasmids, there are three different forms either the triple C forms, OC forms and super coiled form.

So, the transformation efficiency depends on the surface area of the molecule. So, surface area of the molecule the super coiled is smaller or smallest, whereas the other two forms are actually having the larger surface area. So, that is why if you use the super coiled DNA, the transformation is actually going to be very high compared to the closed circular DNA or the open circular DNA. Similarly, the plasmid size if you are working till 10 kB, the transformation efficiency is going to be very high. But if you go beyond that, if you go above to the 10 kB fragments, the transformation efficiency is going to be very, very low.

And in those cases, you might have to use the electroporation or the other methods, you cannot rely on the chemical methods. Then the genotype of the cell, cloning strains for example, E.coli K12 strain have 4 to 5 times the transformation efficiency of the similar strain. For linear DNA, which is poorly transformed in E.coli, the RecB or RecD mutation can significantly improve the efficiency of its transformation.

Then the third is growth of the cell. So, log phase cells are best for preparing the competent cells. Then the method of transformation. So, we already have discussed that the chemical methods are actually going to give you the less transformation efficiency compared to the electroporations. And then we have to also have the damage. So, exposure the DNA to UV radiation is actually going to be another factor which is all to affect the transformation efficiency.

So, this is all about the transformation of the E.coli cells. And I am sure you could have got the better practical experience of how to perform the competent cells and how to you know, perform the transformations. Now, let us move on to the next step. And the next step is transformation in the yeast. So, yeast is also another host which you can use for the overexpression purpose.

And there are many methods which are available in the yeast. The worst method is lead acetate single standard DNA PEG method. So, in this method the yeast cells are incubated with a transformation mixture of the lithium acetate PEG 3500, PEG-1, polyethylene glycol PEG. So, PEG is polyethylene glycol, polyethylene glycol. Single standard carrier DNA and the foreign DNA or the foreign plasmid that is the recombinant DNA at 42 degrees Celsius for 40 minutes. The purpose of adding the carrier DNA is to block the non-specific site on the cell wall and made the plasmid available for the uptake.

Post transformation the cells are pelleted to remove the transformation mixture and re-suspended in 1 ml of water. It is plated onto a solid media with the appropriate selection pressure such as antibiotics. The second method is called as the isperioplast transformation method. An isperioplast transformation method has multiple steps. So, in this method the yeast cell wall is removed partially to produce the isperioplast and that you are going to do with the help of enzyme which is called as zymolase.

So, isperioplasts are very fragile for osmotic shock but are competent to take up the free DNA at a very high rate. In addition, the polyethylene glycol PEG is used to facilitate the deposition of the plasmid and the carrier DNA on the cell wall for the easier uptake. So, how you are going to perform the transformation of the yeast with the help of the isperioplast transformation method? So, in the step 1, in the isperioplast method the yeast cells are incubated with the zymolase the enzyme. So, first you are going to grow the cells at log phase and then the step 2 you are going to incubate the cells with the enzyme which is called as zymolase. So, what the zymolase is going to do is it is actually going to chew up the cell wall at a very discrete step.

So, if you are not going to do a complete hydrolysis or even complete removal of the cell wall you are going to make the partial removal of the cell wall to produce the isperioplast. So, the species or the cell what is going to be generated once you have removed the partial cell wall then it is called as isperioplast. Now, what you are going to do is you are going the step by step these isperioplast are going to be collected by the centrifugation and incubated with the carrier DNA and the plasmid DNA for 10 minutes at room temperature. So, take the isperioplast you collect the isperioplast by centrifugation and then you are incubating the isperioplast with the PEG carrier DNA and as well as the recombinant plasmids. Then in the step 3 it is now treated with the PEG and the calcium for 10 minutes with the gentle shaking.

So, you are going to know keep the cells like the isperioplast the carrier DNA and plasmid and then you are slowly slowly slowly you are going to add the PEG and as well as the calcium. The transformed isperioplast are plated onto a selective solid media and incubated onto the 30 degree Celsius for 10 days because it takes that much amount of time for getting the transformations. So, once you have got this you know transformed isperioplast you are going to plate them onto a selective media and then you incubate that into a 30 degree Celsius incubator remember that the yeast is growing optimally in 30 degree compared to the 37 degree Celsius and then you incubate that 4 days after 4 days you are going to get the plate with colonies and all these colonies are a distance for that particular antibiotic. So, this is all about these the methods the transformation methods what you can use to deliver the DNA into the bacteria or the yeast. In this particular lecture what we have discussed we have discussed about how you can be able to make the use make how you can be able to use the different chemical agents to prepare the you know chemically competent cells and how you can be able to use them for the transforming the with the transformations.

Apart from that we have also discussed about the electroporations and so with this I would like to conclude our lecture here in our subsequent lecture we are going to discuss about the DNA delivery into the mammalian cells. Thank you. Thank you.

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 10

Molecular Cloning

Lecture-44 DNA Delivery (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from Department of Bioscience and bioengineering IIT Guwahati. So, what we have discussed in the previous lecture is that you are going to isolate a gene or the gene fragment from the genome either you are going to aware of the genome sequences or you are not aware of the genome sequences. In both of these cases either you will use the genomic library or the cDNA library or you will use the psi directed primers and you will actually going to use the PCR to get the gene fragment. Irrespective of the sources you are going to get the gene fragment.

This gene fragment has to be digested with the restriction enzyme and that is how you are going to get the sticky ends. So, for example in this case if you are using the two restriction enzyme then you are going to have the two sticky ends and the same way you are actually going to treat the vectors. So, in this case the plasmid which is going to be digested with the restriction enzyme RE1 and RE2. So, that is how you are actually going to have the sticky ends like the RE1 and RE2 and once you put them together and you are going to put a ligation reaction with the help of the enzyme T4 DNA ligase you are going to get the chimeric DNA or the chimeric plasmids.

This chimeric plasmid has to be delivered into the host and if you recall in the previous lecture we have discussed about the DNA delivery methods utilizing the transformations. So, we have discussed about the transformations in the bacteria and or we have discussed about the transformation in yeast and in both of these methods you are going to treat the cells with a chemical agent and with the help of the chemical agent it is actually going to change the surface chemistry and it is also going to make the cells competent enough to take up the exogenous DNA and once the DNA is been taken up you are going to put them into the recovery phase and that is how you are going to get the transform colonies. Now in today's lecture we are going to discuss about the how you can be able to deliver the DNA in mammalian cells. So, as we discussed the DNA delivery in host is very important for the protein production and as you can recall when we were discussing about the surface chemistry or the other kinds of proper modulations which are allowing the cells to take up the DNA. But that kind of modifications are not possible in the mammalian system because first of all the mammalian cell wall or does not have the cell wall and the mammalian system is the surface chemistry is very very complicated compared to the bacterial system.

So, in those cases we have the alternate approaches to deliver the DNA into the host. So, we have the four approaches what we can use in the delivering the DNA into the mammalian cells. The first approach is called as the chemical transfection method, the second approach is called as the liposome or the lipoplex methods, the third approach is called as the bacteriophactin and the fourth method is called as the transduction. So, in the first method you are going to use the transfection agents or you are going to use the chemical agents. These chemical agents are actually going to make the complex with DNA in such a way that DNA is going to be taken up by the mammalian cells.

So, they will be going to taken up by the mammalian cells. So, in the liposome or the lipoplex method you are going to do the same thing but instead of using the chemical agents you are going to use the lipids and in this case sometime you are going to use the cationic lipids and in that case the cationic lipids are actually going to bind the DNA and that is how you are going to have you are going to get the DNA lipid complex and this DNA lipid complex can be delivered or can be easily readily been taken up by the cells. In some cases people are also trying with the liposome method. So, where you are actually entrapping the DNA into the liposome and that is how you are actually these liposomes are going to be having the DNA inside and that is how it they are going to be taken up by the plasmids by the mammalian cells. In the bacteriophactin you are using the bacteria as a source to deliver the DNA.

So, bacteria as a source to deliver the DNA. Similarly, in the transduction you are going to use the virus as a source to deliver the DNA into mammalian cells. So, let us start discussing with the first method and that is called as the chemical transduction method. So, in the chemical transduction method you are going to treat the principle. The principle behind the chemical transduction technique is that you coat or complex the DNA with a polymeric compound to a reasonable size precipitate.

It facilitate the interaction of the precipitate with the plasma membrane and uptake through the endocytosis. There are multiple chemical compound have been discovered which can be able to take make the complex and deliver the DNA into the mammalian cells. So, in a chemical transduction method because you know that the DNA is negatively charged. So, you can add the chemical agent which are positively charged. So, once you are actually going to add the positively charged they are going to make the complex with DNA and as a result they are actually going to make the visible precipitate and these visible precipitate are actually going to be taken up by the cell with a process which is called as the endocytosis.

So, these particles will you know going go and sit onto the plasma membrane and then the plasma membrane is going to be taken up inside and by a process which is called as the endocytosis. There are many methods many chemicals what you can use. So, one of the popular method is the calcium phosphate method. So, in this method the DNA is mixed with

the calcium chloride in a phosphate buffer and incubated for 20 minutes afterwards the transfection mixture is added to the plate in a drop wise fashion. DNA complex DNA calcium phosphate complex forms a precipitate and deposited on the cell in a uniform layer.

The particulate matter is taken up by the endocytotic in the into the internal storage of the cell. The DNA is then escapes from the precipitate and reach to the nucleus through a unknown mechanism. This method suits to the cells growing in a mono layer or in a suspension, but not for the cell growing in the clumps, but the technique is inconsistent and a successful transfection depends on the DNA phosphate complex particle size and which is very difficult to control. Which means in this case you are actually going to take the DNA and then you are actually going to add the calcium chloride along with the phosphate buffer.

Okay. So, what will happen is the calcium phosphate is going to react and that is how the DNA is going to make a complex with the calcium phosphate and once it forms the complex with the calcium phosphate it is actually going to form the particles like structure. So, it is going to perform the precipitate. So, imagine that you have a cell, you have a single mono layer of cell. Right. So, all these particles are actually going to sit on top of this.

So, once the particle sits on any cell, the cell has an inherent tendency that it is actually going to eat these cells or these particles. Okay. Just like as we take the food particles, for example. So, when we take the food particle and it goes inside the you know our body, the cells are actually going to take up this food.

Okay. And as a result, what will happen is the this particle is going to be cut inside, then once the particles are going to be inside the DNA is going to be released from this and that is how the DNA will actually going to reach to the nucleus. Now, what is the disadvantage? The disadvantage is that if this particle size are small enough, they will all they are not going to cause any damage to the cell, but they are if they are big enough if they are going to grow because more and more calcium phosphate if it reacts with the DNA, the size of this calcium phosphate particle is going to grow up. So, if the size is very high, then it is actually going to cause the damage to the cell. So, disadvantage of the calcium phosphate method is the severe physical damage to the cellular integrity due to the particle matter particulate matter sitting settling onto the cell it results in the reduced cellular viability and the cytotoxicity to the cell. So, one of the major disadvantage of the calcium phosphate method is that it is actually going to give you the very low recovery because if you are if you are very good and you are controlling the events in such a way that you are going to make the particle size very small, then it is going to work if it is dozen, then you know the particle size are going to be very big, it is going to cause a physical damage to the cellular integrity and that is how it is actually going to kill the cells.

So, what is the alternative? The ultimate alternative is that you may go with the polypexis method. So, polypexis method is that in a is an alternative method, which was evolved

where the DNA was complexed with the chemical agent to form the soluble precipitate through the electrostatic interaction with the DNA. A number of polycationic carbohydrates such as DAE dextran, positively charged catenin lipids such as transectin or polyamines etc. are being used for this purpose. The soluble aggregates of the DNA with the polycationic complex is readily being taken up by the cell and reaches to the nucleus for the expression.

So, what you are going to do in a polypex method, we are instead of using the calcium phosphate, you are going to use the material which is going to make a soluble precipitate with the DNA and in this category, you can use the DAE dextran or positively charged lipids or you can use the polyamines. All of these are actually going to make a complex with DNA and these complexes are going to be soluble in nature and that is how they are actually going to be taken up by the cell. So, what you are going to do is you are going to take up the you are going to take this chemical agent polypex method or you are going to take the plasmid in two vials and then you will mix them together. So, the mix the equal volume of transectin and the DNA solution, then you incubate for 20 minutes to form the DNA liposome complex. So, the DNA liposome complex is formed, then you can take the plated cells and then you can actually be able to add this drop wise onto the cells.

So, when you drop add the drop white, it is going to eventually going to spread on this and since it is actually going to be a soluble aggregate, it is going to still form the aggregates, these are going to be taken up by the cell and that is how you are going to get the expression of these cells. So, their DNA will enter into the cell and then the DNA will go to the nucleus for the transfection and as well as translations. Then we have the second method where you are going to use the liposome and lipoplex transfection method. So, the liposome and the lipoplex method, another approach of DNA transfection in the animal cell is to pack the cell in a lipid vesicle or liposome. In this approach, the DNA containing vesicle will be fused with the cellular membrane and deliver the DNA to the target cell.

Depression of the liposome and encapsulating DNA was a crucial step to achieve the good transfection efficiency. Liposome prepared with the cationic or the neutral lipid facilitate the DNA binding to form the complex or the lipoplex and allow the uptake of these pluses complexes by the endocytosis. The lipoplex method was applicable to a wide variety of cells and found to be transfect large size DNA as well. Another advantage of the liposome or lipoplex is that the addition of ligand in the lipid bilayer. It can be used to target a specific organ in the animal or a site within the organ.

So, in the liposome or the lipoplex method, you are going to either use a cationic lipids and make the DNA protein come the lipid complexes or you are going to make the liposomes and you are going to entrap the DNA into that. So, once you prepare the liposome, it is going to go and fuse with the cells and that is how it is actually going to give you the DNA, they deliver the DNA. Now, let us go to the next method. And the next method is the liposome and the lipoplex transfection method that is anyway we have discussed. So, this is the basic principle that you have the lipofectamine agent and that is actually going to make the

complex with the DNA.

So, DNA complex is going to be taken up by the cell by a process which is called as endocytosis. And from the endosome, this DNA is going to be released and this DNA will go to the nucleus for the expression studies. Now, let us go to the next method and the next method is the bactofectin. So, bactofectin is the method where you are going to use the bacteria for the DNA delivery. So, bactofectin is more common in term in the case of plants.

So, this mode of gene transfer is very popular in the plant where agrobacterium tumefaciens is used. So, in animal cells, the bacteria is actively being taken up by the host cell through a process which is called phagocytosis and the entrapped in a membranous vesicle known as phagosome. Then the bacteria get escaped from the phagosome and get lysed to release the DNA into the cytosol. In alternate mechanism, the bacteria get lysed inside the phagosome and the DNA is released into the cytosol. The bacteria species used in methods are *Escherichia coli*, *Salmonella*, *Shigella*, etc.

Most of the strain used to deliver the DNA are attenuated, so they should not harm the host cells. So, in a bactofectin what you are going to do is you are going to take the bacteria. So, this is a bacterium cell and you take the DNA and then you mix them together. So, bacteria will take up this DNA and that is how it is actually going to form the DNA bacterial complex. Once the DNA bacterial complex is formed, it is going to be taken up by the endocytosis or the phagocytosis.

And once it is going to be taken up by the cell entry process, it is going to be present in the membranous vesicle and from this membranous vesicle, the DNA is going to be released and this DNA will go into the nucleus for the expression studies. Then we come to the last process and that process is called as the transduction. So, in the transduction you are going to use the virus as a source to deliver the DNA. So, transduction or the virus mediated DNA delivery into the mammalian cells. So, virus particle has a natural tendency to attack and deliver the DNA into the eukaryotic cell.

Most of the viruses they do not have their own cellular machine treat for replication and that is why they have the inherent tendency that well if you add them to the mammalian cells, they will go and attach to the mammalian cell and that is how they will actually going to inject the DNA into the cell and then this DNA will go directly to the nucleus and it is actually going to recombine with the genome and that is how it is going to be a part of the genome. And then once the genome is going to replicate, it is actually going to make the multiple copies of the virus and that is how your the virus is going to spread throughout the body. So, utilizing or exploiting that mechanism we can actually be able to deliver the DNA into the eukaryotic cell. So, cloning the gene of interest into a viral vector is a innovative way to deliver the DNA into the host cell. If the recombination sequences are available, the delivered DNA is integrated into the host and replicate.

Virus has essential component for expression of protein required for DNA replication, RNA polymerase and the other ligand for the attachment onto the cell. In addition, it has additional structural component to regulate the infection cycle. The viral vectors containing cassette to perform all these functions, then it is fully sufficient to propagate independently. Few virus sustain may cause disease if their propagation will be uncontrolled. A mechanism has been devised to keep a check on the uncontrolled proliferation of the virus in a cell.

Through crucial structural blocks are placed on another helper virus. In this case, the virus propagate only if the helper virus has been supplied along with the viral vector. This particular arrangement is made with the virus stain which can cause a disease after integrating into the genome, such as the lentivirus. So, in the transduction species, you can have the two different types of viruses, adenovirus or you can have the lentivirus. Okay, so adenovirus can be used very extensively to express the protein where the adenovirus is going to attack to the cell, they will deliver the genome gene of your interest to the nucleus and that is how you are going to work.

In the lentivirus, lentivirus are more infectious and self replicating. That is why they infect they will be going to when you use the lentivirus, the lentivirus can be used in two cassette. So cassette 1, which will actually going to have the your gene of interest and the cassette 2, so cassette 1 and then the cassette 2, you can keep all the essential genes. So, once you supply both of them, they will actually going to replicate and they will actually going to supply your gene. But as soon as you do not add the cassette 2 or you remove the cassette 2, it will still be able to deliver the gene, but it will not be able to cause the infection.

Okay. So that is how you can actually be able to control the activity of some of the infectious viral particles such as the lentiviruses. So, in a transaction, what happened is that you are actually going to use the, you know, for example, in this case, we have taken an example of bacteriophage. So, in the step 1, the phage is going to inject the DNA into the host. Then the phage enzyme are going to break down the host DNA. So, it is going to break down the host DNA and the in the step 3 to 4, the cell creates no phage including the phage and the host DNA.

And in the 5 to 6, the transducing phage insert the donor DNA and the donor DNA included in the recipient chromosome due to the recombination. So, in the last step, it is going to recombine and that is how whatever the DNA you have, the phage has injected, it will be a part of the host genome. And that is how once the host is replicating, it is also going to replicate the viral genome. So, this is the different method what we have discussed for delivering the DNA into the host system. Now, in the, what we have discussed, we have discussed about the isolation of the gene.

And we have discussed about the isolation of genes. So, we discussed about the two approaches A and B. One is the genomic library approach and the other is cDNA library approach. And these are the approaches when you can use when the genomic sequences

are not known. Whereas, you can use the c approach where you can use the PCR with the site specific primer and that you can do if your genomic sequences are known.

Then once you got the gene fragment like this gene fragment, you can actually be able to do the cloning of this gene fragment into cloning of the gene fragment into vector. So, in the step two, so this is the step one, this is the step, sorry, this is the step one, this is the step two, that same you are going to do for the vector also. And then in this, and then there is a step three, step three, you are going to put the ligation reactions. So, that is also a part of the cloning reactions. And the step four, you are going to deliver the DNA into the host.

So, deliver the delivery of DNA, deliver the DNA into host. Then the step five, you are going to screen these transformed clones. So, you are going to do the screening of transformed clones. And the step six, you are going to check the overexpression. So, so far what we have discussed, we have discussed about the step one, we have discussed about the step three, we discussed about step three and step four.

Now, in the subsequent lecture, we are going to discuss about the step five and six, where we are going to first discuss how you can be able to screen the transformed clones or the clone where the DNA is being delivered. And, and the lastly, we are going to discuss about how you can be able to use the different method to induce the protein production. So, with this, I would like to conclude my lecture here. In our subsequent lecture, we are going to discuss about the, the screening of the clones and as well as the overexpression. Thank you.

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 10

Molecular Cloning

Lecture-45 Screening of recombinant Clones

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bionining IIT Guwahati. So, in the today's lecture, we are going to discuss about how you can be able to screen the recombinant clone. So, now, let us to talk about the screening. The screening is a very important aspect and when you want to screen anything, right, when you want to screen, you should actually going to screen a population based on the exclusive properties of and when we talk about the vector, the vectors are also or the clone is actually going to give you the exclusive properties and these exclusive properties can be exploited for screening the recombinant clone. So, let us discuss about this in our subsequent slide.

So, screening of the recombinant clone, so you can imagine that this is a vector, right, where you might have cloned. Now, this vector is actually going to provide you the various exclusive properties, so that it can be, it cannot be exhibited by the plane vector, but it can be exhibited by the recombinant DNA. So, one of the screening criteria is that this particular recombinant clone probably could express some enzyme and this enzyme is going to catalyze a reaction where it is actually going to convert the substrate into the product and this product probably could be colored or it could actually be able to give you some green or red or blue, some color and if this product is actually going to give the color to the cell and that is how you can say, okay, blue colored cells are transformed or blue colored cells are actually taken up the DNA of whatever you exogenously added. The second and the most popular method is that you can actually be worked with the antibiotic resistance genes and this antibiotic resistance gene is actually going to provide the survival of the host cell which actually got the DNA.

For example, in this particular vector what you see is it has the ampicillin gene. So, this ampicillin resistance gene is actually going to provide the resistance against the ampicillin and because the plane vector or the plane host will not actually going to survive because it would not actually going to have the ampicillin resistance and that is going to be the criteria what how you can actually be able to use the antibiotic resistance genes. The third is the phenotype, okay. So, phenotype is where you can actually be able to use that for that when you are when the cells are going to you know take the DNA they are actually going to show you some phenotypic changes. So, either of these three broader criteria can be used in different screening methods.

So, the first method what we are going to discuss is the blue white screening and blue white screening is where you are actually going to use an enzyme to convert the substrate into the product and that is how this product is actually going to give the blue colored to the cell, okay. So, it is going to give the blue color to the cell. So, it is actually going to use a chromogenic substrate. The use of the chromogenic substrate to detect a particular enzymatic activity is the basis to screen the desired clone. The most popular system to exploit this feature is called as blue white screening where a colorless substrate is processed to a colored compound, right.

The colorless compound Xgal or it is also called as 5-bromo-4-chloro-3-endoxyl-beta-D-galactosidase is used in this screening method as a substrate for the beta galactosidase. The enzyme beta galactosidase is the product of the lacZ gene of the lac operon. It is a tetrameric protein and it is an initial N-terminal region like the 11241 of the protein is important for the activity of the protein. In this system, the host containing lacZ, lacZ without the initial reagent whereas the vector contain the alpha peptide to complement the defect to form the active enzyme. As a result, if a vector containing alpha peptide will be transformed into the host containing the remaining lacZ, the two fragment will constitute to form the active enzyme.

In addition, the alpha peptide region in the vector contains MCS and as a result of insertion of the gene fragment consequently alpha peptide will not be synthesized to give the fully active beta synthase. The enzyme beta galactosidase oxidizes the Xgal to form the 5-bromo-4-chloro-endoxyl and galactose. The endoxyl derivative is oxidized in air to give a blue colored dibromo-dichloro derivative. Hence, the blue colored colonies indicate the presence of an active enzyme and the absence of insert whereas the colorless colonies indicate presence of an insert. So this is actually going to be reversal of what we have discussed.

So if the enzyme is active it is going to convert the Xgal into the blue colored compound but since we are going to clone the gene of our interest into the alpha region of the protein and once the gene you are going to insert that it will not go to complement the remaining portion which is present in the host and as a result it will not go to show you the activity. So the cells which will not show you the activity and remain colorless are actually going to be the transform cells. So this is what we have explained here. So the beta galactosidase is a protein which is actually going to be expressed for lapoprons. So what you have is you have a LAGG-xene which actually has a missing 11 to 41 region.

So if you express that it is actually going to give you an inactive beta galactosidase whereas this missing region is actually going to be present on to the vector which actually can call as LAGG prime. So this LAGG-xene fragment when it combines with this inactive beta galactosidase it is actually going to give you the active beta galactosidase and this active beta galactosidase is going to convert the X-gal which is a colorless product into a blue

colored product and what reaction it is catalyzing it is actually converting this colorless compound into a blue colored 5,5-dibromo-4,4-dichloroindolyl. So this is the colored compound blue colored compound. So since you are cloning the gene into this particular LAGG-xene area you are even if the fragment is being produced it is also not going to complement the beta galactosidase and that is why you are going to have if you are going to have the two scenarios. In one scenario one when the only the vector is present it is actually going to give you the active beta galactosidase and that is how it is actually going to be able to convert the colorless X-gal into the blue colored product but if the insert is present it is actually going to give you the inactive beta galactosidase and inactive beta galactosidase will not be able to convert the X-gal into the blue colored compound.

And that is why if you see the reaction or if you see the colonies what will happen is that you are going to get the blue colored colonies and you are going to get the colorless colonies. So these colorless colonies are the colonies where you are going to have your recombinant DNA because of simple reason that it is actually going to have the inactivation of LAGG-XE. Then the second criteria is the antibiotic sensitivity. So antibiotics are the drugs which are actually been responsible for inactivation or the killing of the bacteria and it happens because the antibiotics disrupt the some of the you know functioning of the cellular properties. So for example there are antibiotics which are disrupting the translation steps, there are antibiotics which are disrupting the transcriptions and there are other antibiotic which are disrupting the protein synthesis.

So if you want to so if you add the antibiotic into the media it will not allow the propagation of the normal bacterial cells because it is going to disrupt the some of these crucial metabolic pathways. But if you have the antibiotic resistance genes so most of these antibiotic resistance genes are actually going to inactivate the exogenously added antibiotics and that is how it is actually going to allow the proliferation of the bacterial cell if they will actually going to have the transform bacteria. So in this case vector carries a functional selection marker such as the antibiotic resistance genes and to be used to select the clones. The antibiotic resistance gene product has a multiple mechanism to provide the resistance in the host cell. In this approach a circular plasmid containing antibiotic resistance can be able to replicate into the host cell plated onto antibiotic containing media.

In the cloning of a fragment into the plasmid, the plasmid is cut with a restriction enzyme and a fragment is ligated to give circular plasmid with insert. The transformation of the both DNA species cut plasmid and the circularized clone into the host and plated onto the antibiotic containing solid media. These circularized clone will give colonies whereas cut plasmid will not grow as it has lost antibiotic resistance genes. So this is the table what I have given and you will see that these are the antibiotics like the ampicillin, kanamycin, tetracycline, gloramphenicol and these are the gene product from the antibiotic resistance genes. So it will actually going to be beta-lactamase, neomycin, phosphotransferase, ribosomal protection proteins and the chloramphenicol acetyltransferase.

So these are the gene product which are going to be responsible for the inactivation of these antibiotics. And what is the mechanism? For example, in the case of beta-lactamase it is actually going to degrade the ampicillin. Similarly in the neomycin phosphotransferase it is actually going to make the covalent modification of the kanamycin and as a result the kanamycin will not be able to you know do its action. Similarly we can have the ribosomal protection protein which is actually going to have the efflux of the tetracycline outside the bacteria. So that is how since the tetracycline will not be able to enter into the bacteria it will not be able to interfere with the protein synthesis.

Similarly we can have the cat genes and the cat gene is going to acetylate the chloramphenicol to acetylchloramphenicol and that also is going to interfere with the action of the chloramphenicol. So what we are going to do in this is we are going to have the two DNA species. One is this is the recombinant clone where you have cloned the DNA into the MCS and it is going to be circularized. Whereas once you have the cut vector it does not have the you know its circular DNA. So once you transform and if you put it onto the ampicillin containing plate this bacteria is actually going to grow because it has the ampicillin resistance gene and that will actually going to degrade the ampicillin.

So it will actually allow the bacteria to grow whereas when you do the transformation of this cut vector since the cut vector you will not be able to replicate it will actually going to will not be able to express the ampicillin resistance gene and as a result it will not be able to go and form the colonies. Then the third approach is the insertional inactivations. So insertional inactivation in this approach a foreign DNA is cloned within the coding gene responsible for a phenotype. As a result of insertion the gene product is not available to modulate the phenotype of the host. This approach is known as insertional inactivation and it can be used for suitable gene genetic system.

For example, in an insertional inactivation of the lacG right. So lacC is a part of the lac operon and it is responsible for the synthesis of beta galactosidase and you know that the X-gal system can be used to detect the insertional inactivation of the lacZ gene to screen the cloned fragment. If the gene is inserted into the lacZ the clone will not be able to produce a functional beta galactosidase. Hence blue colored colonies indicate the presence of an active enzyme or the absence of insert whereas the colorless colonies indicate the presence of an insert. So this is what you have a lacZ which is actually going to produce the functional beta galactosidase and that is actually going to convert the X-gal into a blue colored compound.

And as a result what you are going to see is you are going to see a blue colored colonies. But if you have the bamtron site and if you use this bamtron site which is there in the lacZ and if you use that and you will insert then you are actually going to put your insert within the coding sequence of the lacZ. And as a result what will happen is that it is actually going to give you the non-functional beta galactosidase and if you have the non-functional beta galactosidase it will not be able to catalyze this particular reaction and as a result it is

actually going to give you the colorless colonies. So these colorless colonies are the colonies which are actually going to have the recombinant DNA. So it is actually going to say that okay recombinant DNA is present.

So if you are transforming the vector and if you are transforming the vector which contains the recombinant DNA the colorless colonies are going to say that it is a vector which contains the recombinant DNA. Then we have the instructional inactivation of the antibiotic resistance genes and that we in this example I have taken from the vector which is called as PBR-322. So in the PBR-322 it has the two antibiotic resistance genes. It has the ampicillin resistance gene and it has the tetracycline resistance gene. So if a gene fragment will be cloned in SCA1 which is a restriction enzyme it will disrupt the ampicillin resistance gene and as a result the clone will be ampicillin sensitive and tetracycline resistance okay.

So whereas the original plasmid will be ampicillin and tetracycline resistance. To select the clone first the transformed E. coli is plated onto a tetracycline containing media. Subsequently a replica plate will be made on the ampicillin containing medium to identify the clone growing on the tetracycline media but not on the ampicillin media which means if you have cloned a fragment into a SCA utilizing the SCA as a restriction site what will happen is that it is actually going to disrupt the ampicillin resistance okay. So it is going to disrupt the beta lactamase gene.

So as a result this particular clone is going to be sensitive for the ampicillin action. Now what we are going to do is we are going to transform this onto the into the bacteria and we are going to get the colonies. Now this transform bacteria what you will do is just make a replica plate of this plate and then you grow this first with the tetracycline okay. So when you grow them with the tetracycline both of these clones are actually going to grow okay and then if you put them onto the ampicillin what will happen is that it is this ampicillin resistance is actually going to kill some of the bacteria. For example in this case if you compare this and that what you see is this particular bacteria is being not present here.

So and similarly you can have some more bacterial colony which are going to be present in the presence of tetracycline but they will not be present in presence of ampicillin. So these are the clone which are actually containing the recombinant DNA okay. So what you are going to do is you can just go back and take out these clones from the master plate and that is how you are going to be able to select the transform plates or you are going to be select the colony which contains the recombinant DNA. Then we have the third example of insertional inactivation of the CI repressor. So CI repressor is a protein which is responsible for the shuttling of the virus between the lytic phase and the lysogenic phase.

So during an infection cycle the virus undergoes a lytic and the lysogenic stages and the CI repressor is a protein which is going to you know function as the shuttling protein. So the lytic cycle phase is responsible for the lysis of the host to release the virus particle whereas the lysogenic phase allow the replication of the virus without of lysis of the. The CI gene

encodes for a CI repressor and which is responsible for the formation of the lysogens okay. In the presence of the functional CI the plaque containing unlicensed host cells and has a turbid appearance whereas in the absence of it, it will be clear. This feature can be used to screen the clone to detect the functional CI or the absence of CI okay.

So if we have a functional CI it will say that you do not have the recombinant DNA if you have the non-functional CI then it will say okay recombinant DNA okay. So this is what it is shown here right. So CI is a temperature sensitive repressor so if you change the temperature the CI repressor is going to be expressed and it is actually going to shuttle the virus from the lytic to the lysogenic phase okay. And if it is in the lysogenic phase it will not go to allow the formation of the plaques. But when you clone the protein and you clone it into the CI repressor gene it is actually going to produce a non-functional CI repressor okay.

And when you present the non-functional CI repressor it will actually go to shift the protein shift the cycle towards the lytic phase okay. And as a result it is actually going to form the plaque which are going to be you know which will say that it is actually going to be the presence of recombinant DNA. Then the third approach is the complementation of the mutations. So complementation of the mutation in this approach a mutant gene can be used to screen the plasmid containing the missing gene and the transcomplement will grow only if the gene product from the clone will complement the function. In general the gene taking part in the metabolic pathway or biosynthetic pathway are routinely used for this purpose.

There are three important requirements in this approach okay. So what is the complementation of the mutation is that the host is mutated for a crucial gene okay. So this host will not grow until this particular gene product. So this it is actually missing with this particular gene. So if you supply the gene product right if you supply the gene product which is present on the vector then this is actually going to grow okay.

So this is called as complementation that the host is mutated in such a way that it will not grow until you provide the gene product and that gene product you are actually going to provide by the recombinant DNA. So in this there are three requirements of this approach. The host strain deficient in a particular gene if the gene belongs to the biosynthetic pathway the mutant host in the case are called auxotroph as host depends on the gene product or the final product of the biosynthetic pathway as a supplement in the media for the growth. So in some cases this particular gene could be a part of the metabolic pathway and gene could be actually be responsible for providing the some crucial biosynthetic molecule. So either you provide the gene product or you can be able to provide that product into the media okay and as a result it is actually going to give you the growth of this mutant or this mutated host in the presence of this particular product in media.

For example if this gene is responsible for the synthesis of uracil okay so if you do not have this gene the uracil will not be able to synthesize and this particular host will not be able to

grow but if you provide the uracil into the media then if you add the uracil into this media then this media is actually going to supply the nutrient and that is how this host is actually going to grow. Then a defined media so you should have a defined media with the missing nutrients right because while you are doing growing this particular host you actually can use the media which actually contains the uracil but you should also have a defined media where this particular nutrient is also missing so that when the nutrient is missing the host is looking for that particular nutrient and that nutrient you will get if the vector is going to supply the gene and then you also require a vector containing gene to supply the gene product to complement. Now let's see how it works so you can actually be able to do the complementation to the mutation in a positive feed selection or you can actually be able to do the negative selection. In the positive selection in the positive selection host strain does not grow on the media lacking a functional gene but the host transform with the recombinant clone can be able to supply the gene product required to grow in the media so that is called as a positive selection. Positive selection means you are supplying the gene product from the recombinant DNA and that's how the host is actually going to survive and it will actually grow and it will give you the colony.

Negative selection negative selection is that when you are actually going to you know restore the activity of the gene it will actually going to kill the transform host. So in the negative selection a chemical compound is added to the media which will be converted into a cytotoxic agent in the presence of the gene product and as a result it does not allow the growth of the wild type but the host strain transform with the recombinant clone has non-functional gene product and it grow in the presence of compound in the media. For example in this particular case we have taken an example where we have taken an example of URA3 so URA3 is a gene which codes for the orotidine 5 prime mono phosphate or OMP decarboxylase and an active enzyme process this particular compound which is called as 5 fluoro orotic acid to a toxic compound which is called as fluorodeoxyuridine and generation of the this toxic compound kills the cells carrying the functional URA3 genes. So what we have is we have the gene of URA3 which actually provides a pro enzyme which is called as OMP decarboxylase and OMP decarboxylase process this particular compound like 5 fluoro orotic acid to the fluorodeoxyuridine which is a toxic compound and when the toxic compound is being generated it will actually going to kill the cells. This means if you have the functional OMP decarboxylase it is actually going to indicate that there is no recombinant DNA.

Similarly if you have cloned the fragment within this particular gene then what you have done is you have done the insertional inactivation of this particular gene. Now if you have done the insertional inactivation or you have produced the non-functional OMP decarboxylase this non-functional OMP decarboxylase is not going to you know convert the 5 fluoro orotic acid to the fluorodeoxyuridine and as a result you can allow the growth of these cells under URA3 minus minus cells. So this is actually going to give tell you that if you got the colonies this means the OMP decarboxylase is inactive. Then these are the methods are more popular in the prokaryotic system. Let's talk about now how you can be

able to screen the clones into the mammalian system.

So screening of the transfected mammalian cells. First method is the reporter gene assay. So in the reporter gene assay system a chimeric construct is produced with an enzyme gene which is cloned in front of the promoter of the gene of interest. The gene reporter gene the gene reporter gene construct contain a eukaryotic promoter and an enzyme for easy readout. The reporter gene construct is transfected into the mammalian cells with a suitable transfection agents. Afterwards the cells are being stimulated with the agent to stimulate the production of transcription factor to bind the promoter and drive the expression of the reporter gene.

A suitable substrate is added to measure the activity of the reporter gene. So this is what you have the promoter which and you also have a gene a reporter gene which is going to express enzyme and this enzyme is going to convert the substrate into the product. And this product readout you can be able to study with the help of the several methods like you can do the fluorescence you can do the luminescence you can also be able to do the UV festivals. So these are the reporter gene construct what you can use for screening the mammalian clothes like the CAT gene, LAGZ, luciferase, 4a and the GFP. And the gene product are CATG chloramphenicolosyltransferase then LAGZ beta galactosidase is for luciferase and 4a is for alkaline phosphatase and GFP is the green fluorescent protein.

And the reaction what you are going to see for catalyzing is that when you have the chloramphenicolosyltransferase it is going to run like chloramphenicol to acetyl chloramphenicol and so on. So in a typical reaction you can actually be able to use like luciferase for example reporter gene system. So luciferase is an enzyme which is present in the abdomen of firefly, 14S spiralis. The enzyme utilizes the duluciferin as a substrate to form the axial luciferin. In the presence of ATP magnesium luciferin is getting converted into the luciferin adenylate involving pyrophosphate cleavage and the transfer of AMP into the luciferin and the luciferin adenylate undergoes oxidative decarboxylation to form the oxyl luciferin and simultaneously there will be emission of light.

The reporter gene construct containing luciferase is transfected into mammalian cell. The cells are washed with PBS and lysed with a lysate buffer take the lysate into the luminometer cube and luciferin substrate is injected to start the reaction and measured immediately in a luminometer. So these are the reactions what your luciferase is going to catalyze and ultimately it is going to produce the light and this light can be measured with the help of the luminometer. So what you are going to do is you are going to first take the look expressing vector, you are going to clone your recombinant DNA into this right and then you are going to do the transfection. So once you got the transfection you are going to have the eukaryotic cell which has this recombinant DNA which contains the look gene in front of the promoter and then you are what you are going to do is you are going to lyse the cells and that is how you are going to have the cell lysate and this cell lysate you can put into the cell plate.

You can take the black plate right and you can take the negative controls, you can take the positive control and so on and then you can just put it into the luminometer right and what luminometer is going to do is it is actually going to give you the signal for the luminescence and that signal is actually a light which is going to come from the activity of the luciferase. The reporter gene construct containing luciferase is transfected into the mammalian cells, the cells are washed with PBS and lyse with a lyse buffer, take the lysate into the luminometer cubate or plate and you add the luciferase substrate and it is injected to start the reaction and measured immediately in a luminometer. The second method is you can actually be able to use some fluorescent protein to look at the transfection of the screening of the transfected mammalian cells. So, you can take the chimeric construct with the GFP protein for example. So, in the live cells the GFP protein is a good choice as the reporter gene to screen to cells containing recombinant protein, fluorocellently tag with the GFP at their C or the N terminus.

The cells receiving recombinant DNA will give green fluorescence and it can be visualized with an inverted fluorescence microscope and it can be analyzed in a flow cytometer to separate the GFP containing cells from the untransfected cells. Flow cytometer analysis the cell based on the shape, size and fluorescence a non fluorescent cell is giving separate peak as compared to the fluorescent label cell and with the help of the flow cytometer. Both of these peak can be collected in a separate tube, besides GFP the other protein what you can use is RFP and YFP and CFP. And so, what you are going to do is you are going to transfect the cells and what you see in the under the inverted microscope that all the cells are showing a green color fluorescence. And if you want be interested to collect these cells what you can do is you can just put into the flow cytometer and flow cytometer is actually going to separate the molecules based on the fluorescence.

So, these are the control untransfected cells, these are the GFP expressing recombinant DNA containing cells and that is how you can actually be able to collect these cells in a separate tube and they will be the recombinant cells or the cells with the recombinant DNA, cells with recombinant DNA. And these cells you can separate out and that is how you can actually be able to use them for subsequent experiments. So, this is what we have discussed so far how you can be able to screen the compounds you can get the suitable clones. Now, the question is if you have done the cloning and you got the suitable clone how you can be able to verify the the integrity or integrity of the how you can be able to verify the clone.

So, you because that is very important right. So, you can actually be able to verify the clone by several method. So, what we have discussed so far we have discussed about the screening of the clone and now what you got is you got the recombinant clone which are present in the LB agar plate. Now, the next question is how you can be able to verify the clone because verification of the clone is very important to that so that you should not misguide because you know and verification of the clone can be done with the help of the DNA sequencing. You can actually be able to sequence the clone to know that whether the

your the gene of your interest is fragment is also present ok. So, the confirmation of the cloned DNA can be done by the DNA sequencing.

Historically, there are two method of DNA sequencing with a similar principle of breaking the DNA either the chemical method or the enzymatic method into the small fragment followed by the separation and analyze them on a high resolution electrophoresis gel. So, in a typical DNA sequencing what you are doing is you are taking the DNA sequence ok and then you are breaking that into the multiple fragments and all these fragments either you are using the chemical method which is called as Maxam Gilbert method or you are using the enzymatic method which is called as Sanger's method. So, either with the help of the chemicals or the enzyme you are you know breaking this DNA sequence into the smaller fragment and all these fragments are then going to be separated in a high resolution polyacrylamide gels and that is how these fragments are going to be and then the signal are going to be analyzed for interpreting the sequences. So, let us first talk about the enzymatic method or the Sanger's method. So, the first method what is been discovered is the Dideoxy chain termination or the Sanger's method.

This method is originally been developed by the Frederick Sanger's in the year of 1977 for which the Frederick Sanger's got the Nobel Prize. In this method a single-stranded DNA is used as a template to synthesize the complementary copy with the help of a polymerase in the presence of nucleotide. The polymerization reaction contains a primer and a nucleotide. So, you can have the three normal nucleotides and a 2 prime, 3 prime di-epsi nucleotide triphosphate. This means it is actually you are going to take the single-stranded DNA and you are going to perform the PCR with the help of the three nucleotides which are going to be normal plus one nucleotide which is 2 prime, 3 prime d d n t p's.

So, what will happen is so you are going to run the multiple reactions of the same for the same DNA. So, but you will actually going to change the d d n p p's. In some cases you are going to take the ATP, in some cases you will take the GTP, in some cases you will take the CTP and in the other case you will take the TTP. This means for every DNA sequence you are going to run the four reactions and in four reactions you will take either the d d d d t p plus you will take the all other remaining three nucleotides. So, that is how you are going to you know make the four different types of reactions.

When the DNA polymerase utilizes di-deoxy nucleotide as a nucleotide it gets incorporated into growing chain, but the chain elongation stops at the d d n p's due to the absence of the 3 prime hydroxyl group. In the typical sequencing reaction you are going to run four different d t m p's are taken into the four separate reaction and analyze onto a high resolution polyacrylamide gel electrophoresis. The ratio of NTP and DTP is adjusted so the chain termination occur at the each position of the base in the template. Now, this is what exactly you are going to do. In the Sanger's protocol what you are going to do is you are going to first take the DNA sequence and you are going to have the terminal sequence then you add the primers and when you add the primer and you are going to have the two

options either you go with the Sanger's protocol or you go with the labeling as well as the termination protocol.

So, in the Sanger's protocol what you are going to do is you are going to add the d t p's and the clino fragments and you are going to you know you are going to label the DNA and then you are going to divide this DNA into the four reactions. In the reaction A or that will called as A reaction you are going to add the d t t a t p and the remaining three NTP's. In the second reactions so this is the reaction number one in the second reaction you are going which is called as T reactions you are going to add d d d t t p but and the rest three nucleotides which are normal nucleotides. Then in the reaction number three you are going to add the or which is called as G reaction which is called as d d g t p and the three g d NTP's and in the fourth reactions you are going to take the C reactions which is called as d d c t p plus the four three remaining NTP's and after every reactions you add you are going to put them for chase which means you allow the DNA to be synthesized. So, in the step one your primer is added and annealed to the three prime of the DNA template the radio labeled ATP is been added to label the primers then the step three the polymerase reaction is divided into the four reactions and in the step four DNA synthesis continue until terminated by the incorporation of the specific d d NTP's either the A, T, G or C and in the step five a chase of polymerase reaction is performed in the presence of high concentration of NTP's to extend all non terminated sequences into the high molecular weight DNA.

This high molecular weight DNA will not enter into the sequencing reaction. In the labeling or the chain termination protocol this is the labeling and the chain termination protocol in the step one a primer is added and annealed to the three prime of the DNA template then the step two a limited amount of NTP's are added along with the one of the radio labeled nucleotide to label the DNA throughout the DNA throughout the length. Then the step three the polymerase reaction is divided into the four reactions just like as we discussed here and the polymerase reaction continues with the four nucleotide and out of four one of them would be the dideoxy NTP's. Synthesis is terminated at the specific d NTP's either A, G, C or T to give the DNA fragments of the different length. Now what you got the DNA fragments of the different length right from see for example, from the same DNA you are going to get first big strand then you are going to get this then you are going to get this like this right and so all these fragment has to be analyzed onto a high resolution SDS page and when you analyze them you are going to get this kind of fragments. So you are going to divide that into four reactions A reactions, T reactions, G reaction and C reaction.

Now imagine that we have started with this particular DNA sequence. So A reaction you are going to get a fragment here right for T reaction you got the fragment here and for G reaction you got this spot here and for the C reaction you got the fragment here. Now what you have to do is you have to run in a reverse orientation like this. You have to walk and you interpret this you are going to walk like this and that is how you are going to get first sub A then AT then ATT then ATT A. So that is what you are going to do and that is how you are going to get the complete fragment DNA or the sequence of the DNA what you have

started

with.

Now the second method is the Maxim-Gilbert method. So this method was discovered by the Maxim and the Gilbert in 1977 which is based on the chemical modification and the subsequent cleavage. In this method a 3 prime or 5 prime radio labeled DNA is treated with a base specific chemical which randomly cleaves the DNA at their specific target nucleotide. These fragments are analyzed on a high resolution polyacrylamide gel and autoradiogram is developed. The fragment with the terminal radio labeled appear as a band in the gel which means what it is going to do is it is going to take a DNA fragment it is going to label on one side with the radioactivity and then it is actually going to treat with the chemicals. These chemicals are specific so they will either go to target the A nucleotide they are either going to target T nucleotide G nucleotide or the C nucleotide and as a result what you are going to get is you are going to get the small fragments of this DNA where the terminal DNA nucleotide you know that so on this side you already have a radioactivity.

But the place where it is actually going to be cleaved is either A or T or G or C this means here also you are going to run the 4 reactions the A reaction, T reaction, G reaction and C reaction and that is how you are going to analyze these reactions and that is how they will go to give you the pattern of the DNA sequence. So, the chemical reactions are performed in two step the base specific reaction and the cleavage reactions. So, in the base specific reaction different base specific reagents are used to modify the target nucleotide. Reaction 1 where you are going to use the dimethyl sulfate DMS and that is going to modify the N7 of the guanine and then open the ring between C8 and C9.

So, that is called as G reaction. Then you have a reaction 2 which is you are going to add the formic acid act as a purine nucleotide. So, it is actually act on purine nucleotide. So, it is actually going to be called as G plus A reaction by attacking on the glycosidic bond. Then reaction 3 it is going to use the hydrazine and that is going to break the ring of the pyrimidine. So, it is going to be called as T plus C reaction because it is not specific for the only C or T it is actually going to attack on the pyrimidine basis.

Then in the reaction 4 where in the presence of salt it breaks the ring of the cytosine and that is called as C reactions. Once you are done with this base specific reactions then you are going to have the cleavage reaction. So, after the base specific reaction the piperidine added which will replace the modified bases and catalyze the cleavage of the phosphatidyl ester bond next to the modified nucleotide. This means at the end what you are going to get you are going to get a pattern like this because here also you have added a radioactivity on to the 5 prime end. So, you are going to have the G reaction G plus A reaction T plus C reaction and the C reactions.

The fragment in the G lane is read as G whereas, the fragment present in G plus A, but absent in G is read as A. Similarly, the fragment in C is read as C whereas, the fragment present in T plus C, but absent in C is read as T. To get the DNA sequence the band with the

lowest molecular weight is read followed by the next band in the 4 lane. For example, this means you are going to start from here and you are going to read like this, but there is a issue. G lane the band is of lowest molecular weight followed by the band in the A lane which means if you have the G and if you also have the same band in the G plus A reaction this means you are not going to read this.

It means this is also going to be considered as G reaction. This means from here you are going to read this and then from here you are going to read this from here you are going to see all these two bands are of the same level this means it is going to be read as C. Then from here you are going to read this as G and so on. So, you will actually go with the you will going to go with the from the lower band to higher band. So, for the G between the G and G plus A you are going to if you have the bond in the same line you are going to read them as G.

Similarly, if you have the T plus C band and C band you are going to read them as C. So, if you have got the two bands which are present both in the T plus C and C then you are going to read that as C other than T. So, same is true for the A plus G and G plus G reactions. So, this is the way you can actually be able to sequence the cloned DNA and that is how you can be able to verify the DNA. So, so far what we have discussed we have discussed about how you can be able to utilize the different types of tools or as well as the features what are present in the where in into the vector and recombinant DNA and those are the which you can use for the screening of the recombinant DNA.

And at the end we have also discussed about the DNA sequencing reactions. So, that you can be able to sequence the cloned DNA and that is how you can be able to verify the clone. So, with this I would like to conclude my lecture here. Thank you. Thank you.

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 10

Molecular Cloning

Lecture-46 Protein Over-expression

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati. So, for what we have discussed, we have discussed about the cloning isolating the gene into from the genome either by utilizing the either by utilizing the PCR or with the help of the isolating the clone from the genomic library or the cDNA library. Subsequent to that we have also discussed how you can be able to clone this fragment into a suitable vector and once you got the clone into the vector, you can be able to deliver this DNA into a suitable host and once you got the transformed colonies, you can be able to utilize them for the screening and as well as subsequent to that for the protein production. So, what we have discussed so far is that you are going to get the transformed bacteria and or the transform host that you are going to put it for the screening and once you got the screen, you got the clone containing host cells.

Now this clone can be used for the protein production or the enzyme production. Now when we talk about the enzyme production, enzyme production will depend on the type of host what you are going to use for the production. So as far as the cloning is concerned, the cloning can be done into two different types of vectors. It can be done into the cloning vector or it can be done into the expression vector.

In some cases when you know that the enzyme or the protein what you are expressing is toxic in nature, so in those cases you do not do the cloning related performance or cloning related operations into a expression vector because then you are going to produce a protein and then eventually it is going to kill the cell. So, that is how you are going to use in a cloning vector. Apart from this cloning vector can also be able to use for studying the mechanism of the transcription, replications and the preparation of the genomic and as well as the cDNA library. Whereas the expression vector is exclusively been used for studying the mechanism of the translation or as well as the enzyme production and ultimately it is going to give you the enzyme of your interest. Now when you talk about the enzyme production, the enzyme production is complicated process and it requires the discrete steps.

So before getting into the different steps what you require in the recombinant with the recombinant DNA we should first understand how the protein production occurs in a

particular cell. So, protein production is a multi step process and these steps you have to follow following steps. In the step 1, the binding of the RNA polymerase to the promoter element to start the transcription to form the messenger RNA. So in the step 1 you are actually going to produce the messenger RNA and then this messenger RNA is going to be utilized in the step 2. So as soon as the messenger RNA is synthesized a translational machinery starts the synthesis of the protein or the enzyme.

So protein synthesis starts usually at a start codon which is called as AUG and it is at the stop codon which are called as UAE, UGA or UAG. In bacteria, transcription and translation occurs simultaneously because there is no nucleus right. So, transcription and translation occurs simultaneously whereas in the eukaryotic system, you are going to have the transcription inside the nucleus and then, The translation is going to be in the outside the cytosol. So, That is why in a eukaryotic-system transcription and translation are not going to be together whereas in the bacterial system, you are going to have the transcription and translation together. So, in the step 1 from the gene, you are going to produce the messenger RNA.

So, RNA polymerase is going to sit on to the promoter region and then, it is actually going to synthesize the messenger RNA which is responsible for this. This messenger RNA is going to be modified you know post translational modification and all that. And then ultimately, the first codon which is the AUG is going to be the place where the ribosome is going to sit and that is how it is actually going to start forming the synthesis of the proteins. So, and then it is going to synthesize the polypeptide. This polypeptide is going to be get folded and that is how you are going to get the folded functional proteins.

You can actually be able to read this so that you can be able to understand the subsequent process what we are going to do inside the host so that and it will actually be able to help. So, I have given you a reference which you can actually be below to understand all of these processes. So, you actually required to understand the transcription and translation events, then only you can be able to modulate and you can be able to understand how we are actually overexpressing the protein in the host system. Now, when we talk about the under the in vitro system, under the in vitro system you have the two different species. One is you have the host cells where you are going which you are going to use as the protein production machinery and then you also going to have the transforming agents which will actually going to use for providing the instructions.

This means you are going to first use the instructions and then you put it that into the production machinery and that is how you are going to use that production machinery for the protein production. So, as far as the host cell is concerned you have the multiple choices. You can use the prokaryotic system either the E. coli or other bacterial cells, you can use the eukaryotic system such as you can use the yeast, animals and plants and as far as the transforming agent is concerned you can use the different types of plasmids, you can use the mammalian vectors, you can use the yeast vectors and that all we have discussed

when we were discussing about how you can be able to clone a particular gene fragment into a vector of your choice. So, what we are going to discuss is what we are just going to discuss about the prokaryotic and thresholding of the eukaryotic expression system we are going to discuss about yeast expression system.

Then we are going to discuss about the animal expression system. So, these are the things we are going to discuss. E. coli as a expression system which is going to be for the prokaryotic system, yeast as the expression system which is be part of eukaryotic system, then we also going to insect cell line as the expression system and the mammalian expression system. Now, before getting into of these choices, so you for a particular gene you have the following choices, you have four choices or even more than that.

The first question comes how you can be able to select and select the particular expression system because every expression system has its positive and negatives. So, the number of factor need to be considered to choose the host expression system suitable for the over expression of a protein. The first factor is that the quantity of the desired protein. If the quantity protein required in a small quantity any host expression system can be suitable for the purpose. If the large quantity of protein is required such a E.

coli or yeast or baculoe expression system might be more suitable than the mammalian expression system because mammalian expression system is going to give you a very small amount of proteins. Then it also depends on the size of the protein, the E. coli expression system is not preferred for a large protein size of the protein, but an E. coli expression system is more suitable for the large size proteins. Then we have the compatibility, compatibility between the source organism and the expression system.

So, in general a close distance between the source organism and the expression system is preferred at may increase the chances of getting the expression of the clone gene and the presence of protein in the soluble fraction. Then we also require we have to see the downstream application. So, this is the most important criteria to choose a host factor system. If the protein production is for generating the antibody any expression system may suit for this purpose, but if the protein is required for activity or for ELISA then a compatible expression system is preferred. So, downstream application which means where you are going to use this particular protein is very, very important criteria to select any of these hosts.

For example, you cannot use the bacterial expression system in case you are going to use the downstream product for developing the vaccine or utilizing them for using for production generation of the antibodies or something. Because then there is a chance that you might actually be able to get some bacterial products and these bacterial products are many time causes the allergic reactions into the patients. So, if you downstream application is actually going to decide what expression system is going to use. So, in the E. coli expression system, so in a typical component of an E.

coli expression system additional structural features are essential for an expression vector. What we have discussed when we were discussing about the cloning vector we said that it should have the origin of replication, it should have a multiple cloning site and so on. But if you want to talk about the expression vector what you require is you require the promoter. So, for a cloning vector what you require? You require the origin of replications, number 1, number 2 you require the multiple cloning site and number 3 you also require a promoter because the promoter is going to decide the protein production. So, promoter this is the upstream sequence to the gene and provide the docking site for the RNA polymerase.

Then you also require the ribosome binding sites. So, ribosome binding site includes the Shinder-Gano sequences and it is a docking site for the assembly of ribosomes. So, you also require the RBS and RBS is a ribosome binding site and ribosome binding site is very important for the binding of the ribosome so that it is actually going to initiate the translation. Then you also require the termination sites. So, it terminate the synthesis of the messenger RNA and then some cases you also require the affinity tag.

So, affinity tag is not essential. These are the component which are essential. The origin of replication, multiple cloning site, promoter. The presence of affinity tag either before or after the gene sequence provide a mean to purify the protein using the affinity chromatography. So, these are we are going to discuss then only you will understand what is mean by the affinity tag and how it is actually makes the life easy for the researchers to purify the enzyme in a bulk quantities.

Now, as far as the promoter is concerned in a prokaryotic system, the promoter is containing the some of the classical features such as minus 35 regions and minus 10 region. So, TATA box is there. So, you have the TATA box and you also have the minus 35 regions. So, sequence at the minus 10 and minus 35 are crucial to facilitate the RNA polymerase and the subsequent determination of the strength of the promoter. So, as good these sequences are like minus 10 region and minus 35 region, it actually going to decide how efficiently the RNA polymerase will go and sit to these region and that is how they are actually going to give you the better transcription.

The nucleotide substitution in this region is severely affecting the turnover number of RNA polymerase binding and the transcription initiation side. Subsequently, a number of promoters are designed for the over expression of the protein in E. coli using a strong or weak promoter to see it is the over expression strategies. So, we have the IPTG inducible promoter. IPTG is stands for the indolel pyros is a synthetic analog of lactose and it has been widely been used for the to the construct different expression vector to express the protein in E.

coli, the different vector contain the lac promoter or its derivatives. You have the three different types of promoters. You have the lac promoters, examples of the plasmid is Puck

series and PGM. Then you also have the Taq promoters. So, Taq is a hybrid promoter where you have the some region of the tryptophan promoter and the lac promoters.

So, it is a hybrid promoter where minus 10 region is from lac UV pipe promoter and it is used with the minus 35 region of the tryptophan promoter. Example is PKK223-3. Then you also have the Tric promoter. So, Tric promoter it is similar to the Taq promoter except that the distance separating the minus 10 and minus 35 region of the promoter is different from the Taq promoter.

The example is Ptric99A. So, either of these plasmids can be used for generating the recombinant DNA and then you can be able to transform that into the suitable host and then you can be able to use this for protein production. Then we have the bacteriophage lambda promoters. So, this promoter keeps a tight control over the protein production. It is regulated by the presence of repressor CLTPs857 to either the repressor repress the transcription or not. CLTP5857 is a temperature sensitive and degraded at high temperature and consequently in a temperature dependent fashion it represses the transcription at low temperature, but not at a high temperature.

This promoter is useful in cases where the protein is toxic in nature. So, then we also have the bacteriophage T7 promoters. So, similar to the bacteriophage PL promoter the T7 promoters is used to design the plasmid with tight control on the protein production. These vectors contain most of the structural blocks from the PVRT22 and the MCS in front of the T7 promoter to drive the transcription of the insert. Hence vector containing the foreign gene in front of the T7 promoter for the expression.

So, T7 host E. coli also needs the modification to shoots the T7 promoter and host E. coli is been transformed with the plasmid which carry the T7 RNA polymerase gene or the T7 RNA polymerase gene is integrated into the bacterial chromosome. In few host strain T7 RNA polymerase is placed under the tight control of IPTG inducible LAC UV5 promoter to tightly control the production of the T7 polymerase. So, either of these promoters first step is that you are going to transform the recombinant DNA into the host and then you are going to do the protein production. How you are going to do the protein production? In the step 1 you are going to do the transformations.

So, what you are going to do is you are going to take the recombinant plasmid and you are going to do the transformation into the suitable bacterial species or bacterial strain and that is how you are going to get the transformed bacteria. And you can use the multiple method of transformation, you can use a calcium chloride method or you can use the electroporation. Then the step 2 you are going to inoculate the single colony into suitable bacterial media such as LB media and you can allow them to grow into a 37 incubator. So, a single colony of the transformed colony is inoculated into a suitable media and it can grow up to a log phase such as the OD is 0.

6 to 0.7. And then what you are going to do is you are going to induce, so you are going to in the third step you are going to induce the bacterial species with the inducer such as IPTG for 3 to 6 hours to produce the proteins. So, you can do like that and then you are going to do a centrifugation or the collection of the bacterial cells. So, in the step 4 you are going to recover the bacterial bacteria and analyze the protein expression. So, bacteria can be recovered from the culture with a brief centrifugation at 8000 to 9000 RPM and analyze on to the SDS page. The detail of the SDS page will be discussed in a future experiment.

The SDS page analysis of a particular expression study is given and it indicates a prominent expression of the target protein in the induced cell as compared to the uninduced cell. So, once the induction is over you can actually be able to do the centrifugation and that will actually go to give you the bacterial pellet. This bacterial pellet can be analyzed for the protein production. So, what you can see here is this is uninduced cells and these are the IPTG induced cells and what you see here is a very prominent band of the protein of your interest. Now there are many factors which are actually going to decide the protein production into the E.

coli expression system. So, factor affecting the protein synthesis in E. coli. So, first factor is the translational efficiency. So, translational efficiency is governed by the composition of the promoter especially the sequence of the Shine-Dargano sequences which enables the binding of the ribosome protein production machinery. In addition, the distance between the Shine-Dargano sequence and the start codon is also important for the efficient translation.

Moreover, secondary structure of the promoter elements also affect the efficiency of the gene expression. Then the step 2 is the growth conditions. Growth media has a drastic effect on the protein production. Either the media component provide the raw material for the synthesis of the amino acid or provide the amino acid for the synthesis of a protein. In addition, the growth media rich with carbon source may provide high cell mass and as a result it will give you the more amount of proteins.

Then the third is the codon usage. So, third is codon usage. Genetic codes are degenerate and there are 61 codes which are available for the 20 amino acids. So, this is the genetic code what you see. And there are 61 codes which codes for the proteins whereas 3 codons which are for the stop codon. So, these are the stop codons what you see and which does not code for any amino acid.

Except these you are going to have the codon which is coding for one or other amino acids. As a result the organism has a preference towards a set of genetic code. Expressing these sequence requires the tRNA to recognize the genetic code. But if the host expression system has no tRNA or low level of particular tRNA then it will either delay the synthesis or stop the synthesis of a particular amino acid. Consequently, either it will produce less protein or the truncated protein.

So, what it mean is that every organism has a preference over using some of the codons. For example, in the case of the phenylalanine you have the two codons UUU, UUC right. It is possible that in E. coli probably the UUU is more preferred codon. So, in that case it will actually going to have the tRNAs only for the UUU.

But if you are taking a protein which is you know which does not have this codon which you have the other codon then either it will take the time for synthesis of these tRNA molecules or it will actually going to truncate the protein synthesis at that stage. So, that is a very very important criteria to select the host as per the codon what is present in your gene. Then we have the expression of the fusion protein in E.

coli. The protein in the E. coli expression system can be expressed as a hybrid protein where the reading frame of the two gene one for the fusing tag and the other one is for the foreign gene RNA frame. So, fusion tag can be placed either at the N-terminus or the C-terminus. So, these are the some of the fusion tag beta galactosidase, MBP, thioredoxin, poly histidine, GST and alkaline phosphatase and you can actually be able to use these are the vectors what you can use for providing the for tag. And what is the advantage for example, if you have beta galactosidase you can use the you can use that for blue white screening and as well as for affinity purification. Mostly these fusion tags are being used for affinity purification so that you can be you know avoid the contaminating protein because the fusion tag will not be present in other protein but only present in your protein.

So, if you pass through this to affinity column only this protein is actually going to bind and the rest protein will not bind and that is how you will get the purification in single band or single step. The advantage of the fusion proteins it is going to make the easy purification you might have seen that the with the help of the affinity purification it can be a single step purification. Sometime the tags are being put so that you can be able to target a protein into a particular compartment. So, fusion protein can be targeted to the different cellular compartment for various regions such as the periplasm targeting sequences will allow the protein to accumulate into the periplasm and hence can help to the easy isolation. It can also modulate the half life of the protein in many cases a fusion tag hides the potential protease site which are present on the foreign protein and enhances its half life.

Then it also increases the solubility keeping the tag at n-terminal direct the protein synthesis and helps in increasing the solubility of the foreign protein. Then we have the how we can be able to remove the fusion tag. So for many biotechnology application a protein is expressed as a fusion tag with n-terminal or c-terminal tag to easily purify the protein. But after the purification the tag need to be removed for the downstream applications such as vaccine or the protein crystallographic studies.

A list of reagent is given. So, these are the reagent what you can use you can use the cyanogen bromide you can use hydroxyl amine, enterokinase, factor 10a, alpha-thrombin,

trypsin and subtilin and they are all mostly the proteases except the cyanogen bromide which is a chemical and that is going to cleave just after the methionine. So if you have a tag and if you have a methionine here, so this is the tag actually and this is your gene of interest. So it is having a methionine in between. So what will happen is if you treat that with the cyanogen bromide it is actually going to cleave and that is how you are going to get two fragments your tag and the gene of your interest. So in many cases it is very very essential especially the places where you are going to use this protein as a drug for example like for example if you are going to use that like insulin for example.

So if you are going to put a affinity tag on the insulin it may actually cause the allergic reactions to the patient that is why this affinity tag has to be removed. How you are going to remove the fusion tag? So what you are going to do is in general the fusion tag junction point has either the protease cutting site or the site is sensitive for the chemical treatment. Treating the fusion protein with the protease or the chemical agent cut the fusion tag to release the target protein. So passing the cleavage mixture allows the binding of the tag into the affinity column whereas the target protein does not bind and comes out in the flow through. Target protein free of fusion tag can be collected and used for the downstream application.

So this is what you have in this particular we are taking an example of the his tag. So you have made a chimeric protein so that is having the his tag on one side and the protein of your interest on the other side and then what you can do is you can have you can actually be able to treat this with the thrombin because it has a thrombin cleavage site in between. So what thrombin is going to do is it is going to remove the his tag and the protein and now what you do is you load this onto a affinity column and as a result what will happen is that his tag will go and bind to the beads whereas the protein of your protein is going to come out into the flow through and that is how you can be able to separate the tag from the protein of your interest. So this is all about the bacteria as the host system for protein production and what we have discussed we have discussed about the different steps what you have to follow which means in the step one you are going to transform the bacterial species, bacteria with your recombinant DNA and we have discussed many methods what you can use for transformations. Number two you are going to do the screening or the selection of the transformed clones.

Number three you are going to grow a single colony you are going to inoculate into the media and that is how you are going to induce and the step four you are going to induce that bacterial culture with the help of the inducer. So in this particular example we have taken an inducer as the IPTG and then once the induction is over for 3 to 4 hours then you are going to collect these cells by centrifugations and then you are going to utilize these cells for analyzing the protein production in the SDS page and once you are sure that the protein is being produced then you can be able to lyse the cells and you can purify the protein for downstream applications. So this is all about the different aspects of the protein

production in *E. coli* as an expression system. So with this I would like to conclude my lecture here. Thank you.

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 11

Genome Editing

Lecture-47 Genome Editing (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from department of bioscience and bioengineering IIT Guwahati. And what we were discussing, we were discussing about the different aspects of the molecules in the course molecular biology. So far what we have discussed, we have discussed about the cell biology, we have discussed about the biochemistry, we have discussed about the details of the different types of biomolecules and so on. And then we have also discussed how these biomolecules are interacting with each other and in that process we have also discussed about the central dogma of molecular biology, we have discussed about these application, transcription and translations. And in the previous couple of modules, we were discussing about the different types of techniques which are related to molecular biology.

So, we have discussed about the polymerase chain reactions, we have discussed about de brotting techniques. Now, in today's lecture we are going to discuss about the application of the molecular biology in context of correcting the genetic diseases. So, what you can see here is that we have the different types of genetic diseases which is been found into the human populations. So, we have the Huntington disease, we have hemophilia, we have Perkin's disease, we have colon cancers, we have so many different types of diseases which are been found into the human populations.

And how the human population or humans are you know the how the humans are managing these diseases, they are been having the different types of therapeutic options. One of the option is that you are actually going to generate the drugs and these drugs are actually going to overcome the deficiency or the other kinds of you know he is going to take care of the detoxifications and other kinds of things. So, see the mechanism of these diseases are very different. In some cases you are actually over expressing a particular factor or in some cases you are not been able to provide a crucial component and that is how you can may or may not be able to achieve the particular you cannot be able to run a particular metabolic pathway. So, either of these cases you are supposed to provide that particular factor which is responsible for the correcting these diseases.

For example, we have thalassemia, we have the muscular dystrophy, we have the cystic fibrosis, Huntington disease and so on. All these are happening because one of the crucial

factor is either been produced in a large quantity or it is not been about available for the for the human body. And the only way you can actually be able to provide this is that if you take this factor and put it into the cloning factor. So, let us see how the some of the diseases and we will take an example of few diseases how these diseases are been managed within the for the for the human welfare. So, we have the HIV infections, we have the leukemia, we have thalassemia, we have the sickle cell anemia and the Parkinson disease.

In the HIV infections we have the responsible mutated gene that is called as the CCR5 which is a chemokine receptor and CCR5 is the receptor which is responsible for the replication of the HIV virus. And the target cell for this is the CD4 positive T cells because CD4 plus positive T cells are the cells which are responsible for the immune function into the human body. And you are actually producing the CCR5 you know the genes and that is how you are using the adenovirus and as well as the messenger RNA as a delivery vehicle. And that is how you are hoping that if we do that you are been able to provide the CCR5 and that is how you can be able to you can be able to overcome the mutations and that is how you can be overcome the HIV infections. The other example is the leukemia.

So, it is a what we are discussing is the acute myeloid leukemia or AML and in that case you are actually having the mutated gene which is present in the CD 123 or T-RAC. So, T-RAC is a T cell receptor alpha chain and the target cell therapy is by the CAR T cells. So, you can be able to generate the chimeric antigen receptor T cells and these you are actually going to have the messenger RNA which are going to be delivered into a particular T cell type and that is how you are actually going to generate the chimeric cells and these chimeric cells are going to cure the leukemia. Then we have the beta cell cemia, theta cell cemia you have the BCL 11A which is called as the B cell lymphoma factor 11A and the here the target cells are CD4 HSC 7 HSCs or the hemopoietic stem cells and you are actually going to use the messenger RNA for delivering the ah the ah the clone factors. Then we have the sickle cell anemia and in the sickle cell anemia you have the ah BCL 11A and that is how you are going to have the ah CD4 positive HSC hemopoietic stem cells.

Then we have the Perkin's end disease. So, in the Perkin's end disease you are actually going to have the LRRK 2 or Perkin or Pink 1 or the DGA 1 these are the some of the proteins genes which are responsible for the Perkin's end disease and here you are actually going to cure the IPCs which are called as the induced pleuroplatine stem cells or ESCs or embryonic stem cells and the mesenchymal stem cells and the virus what you are going to use is the adeno associated viruses or the adenovirus or the messenger RNA. What is the general approach for treating the genetic disorder? So, for example, if this is the patient what is actually ah you know suffering from a particular genetic disease. So, what you are going to do is you are going to collect the desired cell type for example, if we take an example of T cell. So, you have collected the T cell and suppose this patient is ah AML patient right.

So, what you are going to do is you are going to collect the the target C cells and then using

the genetic engineering tools ah whether it is the CRISPR-Cas or TALEN or all those kind of the tools you are actually going to do the ah the corrections into the genome of these cells. You are going to take these cells you are going to correct the ah the errors. So, you are going to change the mutated genes and then you are actually going to do the growing the transform cells into the in vitro patid dishes and then you are actually going to put these mutated genes ah in back into the patient and what will happen is that these transform cells which are actually going to have the correct form of the gene is going to be ah work for this particular patient and ultimately it is actually going to overcome the ah particular deficiency. Now if we want to do this right we have the ah what is the our ah main goal? Main goal is to correct the genetic information right. So, genetic information is going to be as you remember right when we were discussing about the genome the genetic information will be present in the in the form of genome right.

So, ah in this particular case what you are going to do is you are going to take the genome and within the genome also particular gene may be actually be target DNA right. So, this particular gene has and has some problem right has some mutations or it has some kind of abnormalities. So, that gene has to be corrected. Now we have the two different types of approaches what people are using one is called as the ah traditional approach or ah traditional approach and the modern approach right. So, traditional approach ah will depends on the gene sequence right and it is mostly been using the PCR based method right.

So, where you are actually going to clone a particular gene and and you are going to use that where in the modern approach you are actually going to use the ah nucleases and all other kinds of sites. So, here here you are going to use the ring finger nucleases you are going to use the CRISPR cas and so on. So, ah in this particular module what we are going to do is we are going to discuss about the traditional approach we are also going to discuss about the modern approach to understand how you can be able to correct the ah information into the genome or how you can be able to do the genome editing. Remember that ah there are ah there are excellent ah lectures available on genome editing itself. So, genome editing is a very very vast subject that is why I would only going to superficially going to introduce you to the topic and there are ah very ah good ah MOOCs courses there is a MOOCs course called genome editing itself which is available ah into the ah MOOCs platform and I would highly recommend that if you are interested into the ah ah into this particular topic you should go through with this particular ah course and it may actually give you the better understanding.

What we are trying to do here is that we are just going to tell you about what is the basis of doing the traditional approach, what is the basis of doing the modern approach and so on and so it is not going to be an extensive discussion about these ah approaches. So, let us first discuss about the traditional approach. So, traditional approach is mostly being descended on the genome sequences right. So, ah and it also depends on the one of the basic phenomena of the recombinations right. So, recombination is the core of the genetic

engineering.

Recombination is the genetic process involving the breaking and recombining of the DNA segment resulting in the creation of the new combination of the alleles. This mechanism operate at the gene level fostering the genetic diversity that mirrors the variation in the DNA sequences against the organisms. So, the recombination is one of the basic phenomena through which the ah two particular DNA are recombining with each other and that is how in that process they are sharing or they are exchanging the DNA content and that that mixing mechanism is always being used in the traditional approaches to generate the ah to generate or to correct the particular type of genetic disorders. So, as far as the genetic recombination is concerned recombination could be of the homologous recombination or the non homologous recombinations. Within the homologous recombinations you can actually be able to have the bacterial recombinations or the eukaryotic recombinations.

Whereas, in the non homologous recombinations you have the multiple approaches or multiple options right you have the CSSR, you have the within the CSSR you have the SIR type or you are actually going to have the ditarosin type. Then we have the NHEJ, we have the transpositions, then VDJ actually recombinations. In the within the transposition we have the DNA transpositions, we have the retroviruses, we have the poly A retroviruses. So, let us first start with the homologous recombinations and we will understand first the mechanism of the homologous recombinations and then we will see how the homologous recombination can be able to ah produce the genetic recombinations and how it get that in turn will actually going to change the genome. So, homologous recombinations ah homologous recombination is a genetic recombination process characterized by the exchange of the genetic information between the two closely related or identical molecule of the nucleic acid.

These molecules can be either double stranded or the single stranded and typically composed of the DNA in the cellular organism although in the viruses RNA may be also involved in the recombination process. Mostly it is involved into the repair mechanism and the meiosis of the prokaryotic and the eukaryotes. So you can understand that you have the duplicated parental chromosome. So, this is the two different types of chromosomes which are homologous or homologous chromosomes. So, you can imagine that if this is the chromosome from the male and if this is the chromosome from the female then what will happen is that these homologous recombination the homologous chromosomes are actually going to go through with the process of recombinations and they will actually going to because they are actually having a particular DNA sequence which is they are common right and because of that it is actually going to do the recombinations and as a result of this recombination there will be a genetic exchange of material between the two chromosome during the process of crossing over.

Remember that when we were discussing about the mitosis and meiosis which said that the crossing over is one of the phenomena which is responsible for the genetic diversity

because during the crossing over one chromosome may you know one part of the chromosome may go to the other chromosome and the other chromosome may also change the content right. So, it is actually been responsible for the diverg, genetic diversity and the same phenomena can be used for producing the genetic modifications. So, once these there will be a crossing over between the chromosomes you will see that the pink portion is been given to the to the blue portion and the blue is going to be given to the portion. So, this portion is actually a part of this chromosome and this portion is actually been a part of this chromosome and that is how these portion is now altered in both the chromosome and as a result what will happen is that you are actually going to have the separate features of the offspring. So, and this is just a simple chromosome this you can actually have the genetic recombinations even in this region you can have a genetic recombination in this region you can actually have the you know the.

So, depending upon what kind of wherever you have the sequence similarity it is actually going to do the crossing over. Now in a homologous recombinations with you can have the two different types of pathways within the prokaryotic and as well as the eukaryotic pathway. So, within the prokaryotes you can have the RecBCD pathway and RecF pathway whereas, in the eukaryotes you can have the DSR pathway and as well as the SDSR pathway. So, first discuss about the prokaryotic recombination homologous recombination into the prokaryotic system and then we will discuss about the homologous recombination into the eukaryotic system. So, in the mechanism of the prokaryotic homologous recombinations.

So, this is the DNA of the homologous recombination where you are actually going to have the different types of factors. So, in the first is the initiation. So, RecBCD bind to the blunt or the nearly blunt area of the double standard DNA BRCK. Then it is actually going to have the unzipping. So, RecB and RecC helicase work together to unzip the DNA.

So, this is actually going to happen here right. So, RecB and RecC is actually going to unzip the DNA which means they are actually going to have the helicase activity they will unwind the DNA. Then the RecB nucleates domain cuts the emerging single strand. So, it is actually going. So, then you are going to have the RecB which is also have a nucleus activity and it is actually going to cut one of the strands.

Then we have the chi side encounter. So, unzipping continue until the encountering a chi side and chi side is having a sequence which is called as GCT GGC GG. And then we have the chi side recognitions. So, DNA unwinding pause briefly and then resume at a reduced rate. RecBCD cut the DNA strand with chi and ultimately there will be multiple RecA proteins are loaded onto the single standard DNA with the newly generated 3 prime end and then there will be a homologous search.

So, this is the strand what is going to be available for making the homologous recombinations. And then when RecA coated nucleoprotein filament search for the similar

DNA sequence on a homologous chromosomes, the search induce stretching of the DNA duplex and then the strand invasions. So, the nucleoprotein filament moves into the homologous recipient DNA duplex forming a dual day loop and then we have the resolution options. So, if the day loop is cut further strand swapping forms a holiday functions resolution by the RuB, Abc or Xg can produce the two recombinant DNA molecules with the recipient reciprocal genetic type with the interaction the DNA molecule differ genetically. And alternatively the invading 3 prime end chi can initiate the DNA synthesis forming a replication form and this type of resolution produce only one type of non reciprocal recombinations.

So, this is what exactly it is shown here right you have the RecBZ system it is actually going to unwind the DNA and then it is actually going to truncate one of the strands and then it will reach to the chi side and from the chi side the DNA is going to be Breck and then this is the Breck DNA is going to be coated with the RekA gene, RekA proteins right and then it is actually going to participate into the recombinations with the closely related sequences and as a result of this recombination it is actually going to produce a two different types of DNA with the recombinations. Then we have the another pathway which is called as RecF pathway. So, in the RecF pathway homologic recombination bacteria employ a repair mechanism for the single standard gaps in the DNA when mutation inactivate the RecBCD pathway and additional mutation disable the SCS, SCCD and XO1 nucleus the RecF pathway can also repair the double standard DNA breaks. So, in this you are going to have the initiation and so RecQ helicase unwinds the DNA then RecJ nucleus degrade the strand with the 5 prime end leaving the strand with the 3 prime end intact then we have a RekA binding. So, RekA protein bind to the strand with the 3 prime end and then RecF and RecO and RecR protein aid or stabilize RekA in this process and then there will be a strand invasions.

So, RekA nucleoprotein filament search for the homologous DNA and the exchange place within the identical or nearly identical strands in the homologous strand the strand invasions and then there will be a branch migrations. So, similar to the RecBCD pathway involves the moment of the holiday function in one direction and then there will be a resolution. So, similar to the RecB pathway holiday junctions are cleaved apart by the enzyme in a process of resolutions and both the pathways may undergo alternate non reciprocal type of resolutions. Despite differences in the protein and specific mechanism in their initial phase both the RecBCD and RecF pathway share the similarities. They both require single standard DNA with the 3 prime end and RekA protein for strand invasions additionally the pathway exhibit similarity in the phases of branch migration and the resolution of holiday functions.

So, this is the pathway of the other pathway and the mechanism of this pathway is also going to be same. And then we have the mechanism of the eukaryotic homologous recombination for repair. So, remember that the homologous recombination mechanism is being used only to repair the you know the damaged DNA because the damaged DNA

information is missing. So, you can actually be able to bring that information from the neighboring residues right. So, with the help of the neighboring residues you search for the homologous DNA and then you are actually going to copy the damaged DNA.

So, in the normal cell you are going to have the double standard DNA breaks and that is how you are going to have the activation of the homologous recombinations in the eukaryotic system. So, ATM recognizes the DSB phosphorylate H2X and facilitate the MDC binding and then we have the MRN the MRE1 plus RAD50 and NBS complex is localized to the DSB. Then we have the CT IP creates C prime overhang where its exonuclease activity and then the RPA binds to the C prime overhand. Then we have RAD51 BRAC BRCA1, BRCA2 replace RPA to form the filament of DNA to proceed to the homologous recombinations. So, this is what exactly the you have a double standard DNA breaks and that activates the ATM and then ATM is recognizing the dual standard phosphorylates H2AX and facilitate the MDC binding and ultimately it is actually going to do the the homologous directed repairing of the DNA strands.

Then we have the two different types of proposed pathway for the eukaryotic homologous recombination mediated repairing. You have the classical double standard break repair pathway or the synthesis dependent strand annealing pathway. Both of these pathways are operating in one organism to another organism. So in the classical double standard break pathway you have the three prime invade and intact homologous template, then formation of the double standard holiday functions, junctions and then junction resolution results in the crossing over or the non-crossing over. Similarly, we have the synthesis dependent strand annealing pathway and it is conservative and result oriented exclusively in the non-crossing over events.

So, the key factors and which are involved in their role in the eukaryotic homologous recombination. So, we have the MRN complex and their job is to initially stabilize the double standard breaks. Then we have the ATM or ataxia telestria mutated it is recognizing the double standard breaks. Then we have the BRCA1 and BRCA2 which is a breast cancer associated genes and these are the checkpoint activation and the DNA repairs. Then we have a CTIP C terminal binding protein interacting proteins and it is interact with the BRCA1 and D phosphorylate the 53BP1.

Then we have the RPA1 or application protein A that stabilizes the sub single standard DNA H2AX, H2A stone family member X and that is responsible for the recruitment and the accumulation of the DNA repair protein and then we have a RAD51, RAD51 forms a filament in the onto the DNA strands. Now, if you summarize the homologous recombinations in both the eukaryotic as well as the eukaryotic one what you will see is that the there are proteins or the processes which are involved which are common between the two. So, you have the introduction of the double standard DSB that is not present in the E. coli system, but that is present in the eukaryotic system where you have the SpO1 and HO which are going to be involved into the introduction of the double standard breaks.

Then we have the processing. So, RECBCD complex or the nucleus or helicase and nucleus system and here we have the MRX complex which is involving the RAD50, 58, 60 and nucleases. Then we have the assembly complex formation and filaments, RECBCD or RECF pathway which is going to be RAD52, RECF59 and BRCA2. Then we have the pairing and strand exchange. So, RECA which is involved into the eukaryotic system whereas, here you have the RAD51 and DMC1 exclusive for the muses.

Then we have the branch migrations. So, branch migration here you have the RECUVAB complex whereas, it is unknown into the eukaryotic system. Then we have the resolution of the holiday junctions. So, RECRUVc and here you are actually going to have the RAD51c, XRCC3 complex and WRN and the BLM. So, these are the comparative study of or comparison of the summary of the homologous recombination in the prokaryotic as well as the eukaryotic organisms.

Let us move on to the non-homologous recombinations. So, in a non-homologous recombinations, the recombination involves the physical exchange of DNA segment between the chromosome or the DNA molecule. When this exchange occur between the stretches of DNA with no extensive sequence homology, it is termed as the non-homologous recombination. Unlike homologous recombination, it does not require a double-stranded break in the DNA for the initiation. It is also relatively less precise and error prone and it is often lead to the insertion or the deletion of the nucleotide at the site of recombinations. NHEJ is a DNA repair mechanism that involves the direct ligation on the broken end and thus do not require a homologous template strength.

Whereas in the transpositions, the CSSR or VDJ recombinations and the phenomena of transposition also does not require the extensive sequence similarity between the strands of DNA involved into the recombination process. So, the non-homologous recombination does not require the sequence similarity. It actually happens abruptly and it does not require the double-stranded breaks also for initiation. So, basically the non-homologous recombination is going to be an exchange of the DNA between the you know or with no with no sequence extensive sequence similarity. There could be sequence similarity, but that would be very very minor and in these kind of cases you are actually going to have the multiple examples like transpositions, VDJ recombination which is responsible for the generation of the different types of antibodies and so on.

And in all of these examples, there will be no there will be no sequence similarity which is involved. So, we have the site specific recombinations, so which is called as the SSR. So, also called as the conservative site specific recombination or CSSR is a recombination recombination between the two defined sequence element of the DNA. This process is carried out by the protein known as recombinase which brings together the specific ends of the DNA forming the synaptic complex resolving onto the three outcomes. So, you can have the insertion, inversion and deletions and these are the recombination sites right you are

actually going to have the insertion or inversion or the deletions.

Type of recombination in the CSSR, so based on the amino acid sequence homology and the mechanism followed most recombinations can be classified into two types. It can be serine recombinase or the tyrosine recombinase. Serine recombinase cleaves all fold strand involved into the exchange whereas, the tyrosine recombinase is one strand in each site is cleaved leading to the formation of the holiday junctions. So, tyrosine recombinase and serine recombinase the difference between the tyrosine recombinase and the serine recombinase. So, recombinase it is going to be a cry whereas, the recognition site is going to be locside and it is actually going to you utilize for the circularization of the fast P1.

We have the lambda integrase which is going to be at P and B sites and it is the integration of the fast lambda. Similarly, we have XRD CD resolvase complex which is going to be diff in nucleoid or the serine plasmids. It is the resolution of the dimers into the E. coli. Similarly, for the serine recombinase we have the H-in invertase which is a HX site and the inversion of the promoter in the salmola and then we have TA3 and Y sigma resolvase which is a rest site and it is going to be resolution during the replicative transpositions.

Mechanism of the serine recombinase. So, one molecule of recombinase catalyze the cleavage of a single strand thus a total of four molecule of recombinase is required for the recombination of the recombination process. The three hydroxyl group of DNA at the three prime act as a nucleophile to attack the recombinase DNA complex at the five prime to generate the free recombinase and recombine DNA. Slippery hydrophobic part of the top and bottom half of the recombinase dimer rotate by the 180 degrees and that is how this portion will go here and this portion will go here. So, all the four strands are first going to be break by the serine recombinase. So, you are going to actually going to have the four strands and four strands are actually going to have the one strand will going to have the three prime OH the other one will have the five prime OH and so on and then there will be a you know the dimer will rotate.

So, in that case this portion will rotate on this side this portion will rotate on this side and that is how it is actually going to have the recombination and that is how they are actually going to have the exchange of material between the two strands. Then there will be electrostatic interaction stabilized initial and the rotated state of the recombinase and ultimately there will be a exchange of DNA material between the all the four strands. Then we have the mechanism of the tyrosine recombinase. So, recombinase cleaves one strand at one each side of the recombination. So, they are going to cleave this particular side and this is going to cleave the another side.

So, you are going to have the two strands which are actually been participating into the recombination process. The cut strand is exchanged between the DNA molecule. So, this portion will go into this and this portion will go into this and that is how this will now be a part of this molecule and this will going to be part of this molecule. Then we have the

formation of the holiday junctions and then cut at the second strand. In each side strand exchange occur and the resolution of the holiday junction.

So, once you are going to have the cut on one strand. So, you are going to have the cut on this strand and then this strand and then there will be an exchange between the cut text and cut extent will exchange between the two strands and then there will be a formation of holiday junctions and then you are going to have the cut at the other strands right. And then there will be an exchange of the genetic material and that is how you are going to have the recombination. So, you see that here you have the blue and then you are going to have the yellow which is coming from this strand and then you are actually going to have the red instead of the blue right in this particular strand. Similarly for the lower strands those strand is actually going to have the blue and it is going to also have the red on the other side.

Similarly for this one this one has the red on both sides. So, it is going to have red on both sides, but on this side it is actually going to have the instead of red you are going to have the blue and that is how you both the strands are now going to have the new DNA molecules or recombined DNA molecules. What is the biological significance of CSSR? So, it is actually been involved into the DNA repairs, it is also involved in the gene regulations, it is involved in the genome rearrangements, it is also involved in the horizontal gene transfers, development and so in a DNA repair it allows the precise cut and replacement of the damaged DNA strands making the genome integrity in equalize. Whereas in the gene regulations the phage lambda uses a CSR mechanism to switch between the lytic and the lysogenic cycles. Then the genome rearrangement the CSSR causes the gene rearrangement leading to the various combinations like the inversion insertion or deletion of the gene sequence. For instance, switching in the flagella component of flagella in *salmola* is mediated to the CSSR mechanism.

Then we have a horizontal gene transfer. So, bacteria often use this mechanism to achieve the new traits from the other bacteria or mobile genetic element leading to the bacterial adaptation and evolutions. Then we have the development. So, many organism uses the site specific combination method to regulate the tissue specific gene expression and cell differentiation during the development. Then we have also have a tracking of cell lineage during the development of *drosophila* was done using the FLP FRT system. And then CSR is widely used to manipulate and engineer genomes of organism like the ablation of the gene function induction of gene expression at a specific time during development and etcetera.

Then we have another example which is called as VDJ recombination. So, VDJ recombination is responsible for producing of production of the antibodies. So, it is a specialized set of DNA recombination mechanism that impart the enormous diversity in the B cell and T cell receptors. This recombination process occur between the specific site onto the V, D and J segment in the gene for the generation of immunoglobulin and the TCRs. These recombination sites are called recombination signal sequences or RSS.

There are two different types of RSS, 12 base pair RSS and 23 base pair RSS. The combination cannot occur between the same site of the RSS. So, this is a 23 base pair RSS and this is a 12 base pair RSS. And this is the T cell receptor. So, what you can see here is that the V, D and J are recombining with the different types of modules. So, you have the different variation of the V component, you have different components of the J component and so on.

And that is how they are actually recombining with each other to give you the different types of the T cell receptor. So, this is one of the T cell receptor, this is another T cell receptor and so on. Similarly, in the B cell receptor, B cell also will have the different types of V, different types of J and different types of the D components. So, DJ rearranged DNA joined and that is also you will see that VDJ is giving you the different types of antibodies and different types of the antigen binding sites present into the antibody molecules.

Now, another example is the NHEJ. So, non-homologous endjoining. So, NHEJ is a mechanism for repairing the double standard DNA breaks in DNA. Unlike homologous directed repairs which require a homologous template, NHEJ directly ligates a broken end without the need of a template. This makes the NHEJ an efficient process that can operate both in dividing and non-dividing cell. The term non-homologous endjoining was introduced by the Morie and Haber in 1996. NHEJ is often guided by the short homologous DNA sequences known as the micro homologous.

These sequences are typically found in a single standard overhang at the end of the double standard break. When the overhangs are compatible, NHEJ can accurately repair the breaks. However, if the overhangs are not compatible, imprecise repair may occur leading to the loss of the nucleotide. Inaccurate NHEJ repair can result in the loss of genetic material and may lead to the translocation and telomere fusion. These events are considered hallmarks of the tumor cell highlighting the importance of the proper NHEJ function in maintaining the genomic stability.

NHEJ is wide spread and exists in nearly all biological systems. In mammalian cells, it is the predominant pathway for the double standard break pathway repair. However, in the budding yeast like the *Saccharomyces cerevisiae*, homologous recombinations tend to dominate under the common laboratory conditions. When NHEJ is inactive, double standard breaks may be repaired by alternate more error-prone pathway like the microhomology mediated endjoining. In MMEJ, end resection reveals the short microhomologous on either side of the break guiding the repair. Unlike classical NHEJ, MMEJ often results in the deletion of the DNA sequence between the microhomologies.

What is the mechanism of the NHEJ? So, you have a double standard break and then this double standard breaks are actually going to have the binding of the Ku70/80. So, recognition of DNA ends by the Ku70/80 heterodimers. So, they will go and bind to the DNA

breaking sites and then they will recruit the DNA PKCs then nucleus such as artemis stream the incompatible strands and then the XRCC4 DNA ligase ligate the complete seal of the break and then the DNA double standard breaks are going to be repaired. Key players in the NHEJ and their mutation related diseases. So, you are going to have the different types of diseases which are been found if there will be a mutation of the important components.

You can have the Ku70, DNA PKCs, artemisis, DNA pol and the ligase and XRCC4 and XLF. So, there if there will be mutation of XR70 it will actually be responsible for SID and lymphoma and reduce sensitivity. If there will be mutation in the ligases then it is going to be for immunodeficiency and reduce growth and developmental issues, microcephaly and malignancy. Then if there will be a problem of XLF and other kinds of things then it is going to have the embryonic liability, syndrome, immunodeficiency, developmental delay and microcephaly. Then we will just move on to the next example and the next example is called as

the
transposons.

So, transposons are also been called as the jumping genes right. So, transposons are segment of DNA that then can move from one locus to a relatively nonspecific site in the genome. Movement of transposon can occur without the duplication of the element accordingly they are classified as the class 1 or replicative DNA or the type 2 which is non replicative DNA. DNA transposons carry inverted repeat sequences at their end as recombination site flanking recombination recombinase protein called transposon which bring out the process of recombination. Transposons retrotransposons contains LDR and the gene for integrase and reverse transcriptase activities. So, there are different types of transposons you can have the class 1 transposon which is called as copy and paste whereas, the class 2 transposons which are called as the cut and paste.

So, in the class 1 you are actually going to have the LTR transposons retrotransposons and the non LTR retrotransposons. So, when they will going to do a transcription they are going to produce the messenger RNA and then there will be a reverse transcriptase to produce a double standard DNA and then these double standard DNA are actually going to integrate into the target sites. Similarly for the non LTR retrotransposons there will be a transcription it is going to produce the messenger RNA then there will be target primed reverse transcriptase and it is actually going to integrate into the target DNA. Similarly for the class 2 cut and paste you are going to have the DNA transposons which is going to be excised from the other DNA and then it is actually going to integrate as such into the target DNA. There are examples of transposons you have the bacterial replicative transposon which is called as TN 3 and Faz mu then you have a non replicative transposons like IS element and simple transposons.

So, IS 1, IS 2, IS 50, IS 10, TN 1, TN 7, TN 501 then we have a composite transposon which is called as TN 5, TN 9, and TN 10 then we have a eukaryotic elements like HAT family members and so on. Then we have a virus like transposons like tie element of the yeast and the copia of the drosophila then we have a poly A retrotransposons. So, that is the line and

sign in the mammals or the alu in the humans. What is the outcome of the transpositions? So, altered gene expression through insertion within the gene for example, the ISI element can insert itself nearly 18 inch promoter sequence.

IS element can cause the polar mutations and the reg dependent recombinations. Site specific combination between the transposable element in the same molecule can cause the rearrangements such as the deletion, inversion etcetera and then IS 1 element can cause the deletions and TN 3 and resolve ways can cause the resolve ways reaction towards the deletion and the mutation in general through the insertions. Now these are the some of the approaches which are been very very common or which are been found when the you are using the traditional method of generating the genetic recombinations. So, you can have the homologous recombination, you have a non homologous recombination and so on. Now when you are going to use these methods and you are trying to generate the genetic modifications, you are actually going to face the lot of limitations and you are going to have the lot of issues. What are these issues? These issues are that there will be a lack of precision, there will be a low efficiency, there will be a time consuming and the limited edition scopes and the off target effects.

These are very very serious effects right. You are going to have the lack of precision which means if suppose you use the non homologous recombination methods and so non homologous recombination methods are going to be random. They can be even to one side or they could be another side. Even for the homologous recombination also you require the very precise you know the flanking sequences then only you can be able to edit the particular gene into the genome. But what happen is that even if all these depends on to the sequence similarity. If the sequence similarity is unique then it may actually be able to give you the very precise removal of the mutated genes.

But if the there will be a any you know off targets right if there will be any kind of similarity between the sequences then it may actually replace the another gene or it may actually give you the side effects. Similarly, all these depends on to the integration of the your externally supplied DNA into the system and you can actually be able to use the different types of DNA delivery method to deliver the DNA into the site of actions. That process is very very inefficient and on the other hand once the DNA will enter then it actually should go through with the recombination process and then and during this journey the DNA should not get any kind of DNA's and other kinds of attacks and though also that is why the efficiency of this recombination is also going to be very very low. Third is this all requires the genetic cassettes to be prepared then you are actually going to transform that and so it is actually very very time consuming. And the fourth is because you are dependent on the you know the flanking sequences or you are dependent on the cellular machinery the scope is very very limited.

And on the other hand the first part is that majority of the time when you are you know transforming that cells with the these externally supplied DNA they may actually get

integrate into the off target sites and because of that it may not give you the desirable results. So these are some of the challenges what people are facing when they are using the traditional method for the genome editing. So what we have discussed we have discussed about the homologous recombination we have discussed about the non homologous methods within the homologous recombination we have discussed about the methods in the prokaryotic as well as eukaryotic system and within the non homologous recombination we have taken examples of transpositions we have discussed about the VDJ recombination and so on. So with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss about the modern approach of the genome editing. Thank you. .

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 11

Genome Editing

Lecture-48 Genome Editing (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from Department of Bioscience and Bioengineering IIT Guwahati and what we were discussing? We were discussing about the genome editing in this particular module and so far what we have discussed in the previous lecture, we have discussed about the traditional approaches where we are using the homologous recombinations or the non-homologous recombination as an approach to modify the genome. Both of these approaches are having the many drawbacks and these drawbacks were you know corrected or will be improved very significantly in the modern approach. So, in today's lecture we are going to discuss about the modern approaches how you can be able to use these approaches to do the genome editing. So, as the name suggests the genome editing involves the alteration of the organism genetic code and I have you if you recall in the previous lecture, we have discussed very significantly that how these you know approaches are being used for curing the different types of diseases. Now, in the what was the major drawback of the previous approach was that in a homologous recombination or the homologous non-homologous recombination that they were having the off targets, they were expensive, they were time consuming and these things can be very much reduced with the modern approaches.

So, we have the you know couple of modern approaches what we can actually be able to use to do the genome editing and in the if you see how the genome editing is been evolved that in the beginning people were using the restriction enzyme for cutting the DNA a particular part of DNA and then using that for you know bringing the modification into the system. Then we have come up with the zinc finger nucleases and in the in the year of 1980s, the zinc finger nucleases were discovered and they were actually been engineered in such a way that they are actually going to be used for genome editing. Then we discuss about the TALENs. So, the TALENs are a structure similar to zinc finger nucleases, but they are been used for transcription activator like effector for DNA binding and they are also been used for DNA editing.

Then the modern approach is the usage of the CRISPR-Cas which I think you might have heard about. So, CRISPR-Cas is the latest approach what people are using for genome editing which involves a different types of protein and other kinds of factors. Then we have the base editing which is also been done with the help of the CRISPR and then we have the

primer editing approach which where you are actually using the Cas9 kind of protein and that is also been used for the genome editing. And then we also have the paste where you are actually going to have the different you know techniques to actually be able to use for the genome editing. So, these are the some of the edit approaches what people have used for genome editing.

What we are going to discuss today is we are going to discuss about the zinc finger nucleases, we are going to discuss about the TALENs and we are also going to discuss very briefly about the Cas-Cas9 method. And if you recall in the previous lecture also when I was discussing about the homologous recombination and the non-homologous recombination I said very clearly that these lectures are only to introduce you the topic. They are not going to be extensively been used for discussing the topic. Because there are very good genome editing MOOCs courses are available on to the NPTEL portal and you can actually be able to use or you can be able to follow them by getting the more detail about these approaches. So, in today's lecture we are discussing about we are going to discuss about the zinc finger nucleases, we are going to discuss about the TALEN and we are also going to discuss about the Cas-Cas.

And all these approaches if you see they are actually been having the you know I have been arranged in such a way that they are actually going to have the time and cost. So, time and cost is going to be more for the zinc finger nucleases and it is going to be less for the CRISPR-Cas. But as far as the feasibility is concerned the feasibility would be less, feasibility will be you know more for the CRISPR-Cas and it is going to be less for the zinc fingers. This means in any case the zinc finger the Cas phases the CRISPR-Cas approach is the best suited approach for genome editing. So, let us first discuss about the zinc finger nucleases, then we will discuss about the TALEN and then lastly we are also going to discuss very briefly about the CRISPR-Cas also.

So, zinc finger zinc finger proteins. So, zinc finger is a compact structural protein motif characterized by the binding of one or more zinc zinc ions contributing to the stabilization of its folded structure. The term zinc finger was first pointed based on the finger like appearances observed in the presumed structure of transcription factor 3a from the African frog Xenopus. However, it has been identified in a diverse area of protein structure within the eukaryotic cell. The discovery of zinc zinc protease is in the Xenopus transcription factor 3a essential for its function was reported in 1983 marking marking the first document instance of the zinc requirement for a zinc regulation protein.

Subsequently, the similar findings were observed in the Cooper factor in the Drosophila and the zinc finger commons commonly manifest as a metal binding domain within the multi domain proteins. So, in a typical zinc finger proteins proposed a model for the TF3 proteins 30 kDa region featuring the 9 repeating region of 30 amino acid region the 25 of the 30 amino acid in the repeat fold around a zinc ion to form a small independent structure domain thus finger. And the 5 intervening amino acid provide the linkage

between the consecutive fingers. Each unit forms a zinc coordinate coordinated fingers stabilized by the conserved amino acid and a potential hydrophobic cluster. The rich the cluster the region is rich in basic and the polar residues are implicated in the nucleic acid binding.

So, zinc finger motifs or zinc finger nucleus is a small protein and this is a protein which has been designed where you are actually going to have the zinc finger motif in a you know 30 amino acid residue. And these are been repeated multiple times so that it can be actually have the binding of the zinc to these regions right. So, you see that in a particular one 30 nucleotide 30 amino acid residue you are actually going to have a zinc in the center and it is been coordinated by the many of the residue. Mostly the histidine what is been present on to the different helix and beta sheets which are actually going to be utilized for stabilizing the zinc what is present in the this particular motif. And this zinc particular motif is actually going to have the affinity for the DNA binding.

So, they are been actually been utilized or can be exploited for the genome editing applications. So, folding of the zinc finger relies on the tetrahedral coordination of the zinc ion with the two invariant pair of the cysteine and histidine forming a conserved cysteine-cysteine histidine-cysteine pattern. So, within this you are actually going to have a zinc in the center. Additionally each repeated includes the tyrosine or phenylalanine or phenylalanine 17 and elsin 23 large hydrophobic residues that may interact to create a stabilizing hydrophobic cluster into the compact finger module. The 30 amino acid repeat rich in basic and polar residues concentrated at significant number between the second cysteine and the first histidine emphasizing its role into the nucleic acid binding.

So, what you have here is that you are actually going to see that there is a zinc which is coordinated between the cysteine and tyrosine and histidine residues. In some cases you are actually not going to have the histidine you are also going to have the phenylalanine and the other hydrophobic residues like the leucine and this is kind of going to be an environment which is required for the zinc binding. The first zinc finger was identified in the 1980 for the transcription factor third A which gives rise to the discovery of new group of transcriptional activator protein with the 30 amino acid repeating region. This new class of protein was able to bind the specific sequence of the DNA. So, that is very very important that zinc finger motif or zinc finger is actually going to be utilized for recognizing the specific DNA sequence and that can be modified or that can be modulated in such a way that you can actually be able to use them for probing a particular DNA sequence.

The zinc finger structure is maintained by the zinc ion which coordinated cysteine and histidine in different combinations. Originally the zinc finger referred to a DNA binding motif in a xenophase, but now it includes various structural coordinated coordinating a zinc ion. Initially the classified by the number and order of the histidine residues recent classification focused on the fold groups determined by the overall shape of the folded domain. Common fold group like include the cysteine, cysteine like classical zinc finger and

the zinc ribbon. So, cysteine histidine to zinc finger motif consist of an alpha helix and anti parallel beta sheets the zinc ion is coordinated by the two histidine residues and the two cysteine residues.

That is what I have shown before also that you are going to have a zinc residue which is coordinated between the two histidine molecules and the two cysteine molecules. The cysteine histidine like fold group is by far the best characterized class of zinc finger and it is commonly common in mammalian transcription factor such domain adopt a simple beta beta alpha fold and have the amino acid sequence motif. So, in the amino acid sequence motif you are going to have the X 2, cysteine X 2, then cysteine X 12, histidine and the histidine. The specific amino acid interaction between the zinc finger and the DNA was elusive until 1991. A pivotal movement occur when the scientist successfully deciphered the crystal structure of complex involving a DNA oligonucleotide and the three finger zinc DNA binding domain of the mouse transcription factor ZIF 268.

In this structure the alpha helix play a primary role binding to the DNA major group through the precise hydrogen bonding interaction from the amino acid at a helical portion 1, 3 and 6 to 3 consecutive bases on the strand of the DNA. Despite this strand configuration there are notable variation in the sequences. Consequently the zinc finger motifs are considered excellent natural building block for the DNA for the design of the protein tailored to recognize a specific DNA sequence owing to their adaptability and versatility. So, that is a very very important portion that because of this kind of zinc finger coordination between the histidine and cysteine it is actually going to be very useful in you know exploiting to detect a specific DNA sequence. So, far there was no protein available which actually recognizes a specific DNA sequence.

So, and that was very very important that you actually can be able to target a specific DNA sequence utilizing this particular type of zinc finger. And once the zinc finger is been discovered the people have also started developing the zinc finger nuclease. So, the zinc finger nuclease are synthetic restriction enzyme created by combining a zinc finger DNA binding domain with the DNA cleavage domain. So, that is what it is actually happening. So, in the in a typical zinc finger nucleus what you have is a DNA binding domain which is from the zinc finger right where you are actually going to have the zinc coordinated into the histidine and cysteine and then you just put actually a DNA cleavage domain.

So, it is like a restriction enzyme which is actually going to chew a particular type of part of the DNA. The engineered zinc finger domain can be customized to recognize a specific target sequence allowing the zinc fingers to pinpoint the unique sequences within the integrate genome. Leveraging the inherent DNA repair mechanism in cells these agents are employed to accurately modify the genetic makeup of more complex organisms. Alongside CRISPR-SCAS9 and the TALEN ZFN stand out as a significant instrument in the relief of the DNA genome editing. So, what you are going to do is you are going to have the two tools actually you are going to have a DNA binding domain you are also going to have a DNA

cleavage domain right and utilizing them you can be able to take out a part of DNA right and once you take out this part of DNA it is actually going to be cured in due course of time and that is how it is actually going to give you the modification of that particular sequence.

So, in a DNA binding domain especially has 3 to 6 zinc finger repeats recognizing 9 to 18 base pair utilizing 3 finger array specific 9 base pair target site can be addressed. Similarly, you can actually be able to have the zinc finger nucleases which are actually going to be having a non specific globase domain typically type 2 restriction enzyme like the FOCI requiring the dimerization for the DNA cleavage. Standard zinc finger nucleases link this domain to the zinc finger domain and for the effective cleavage the two zinc finger nucleates must bind opposite DNA strand with a specific spacing. Protein engineering technique including the direct directed evolution and structure based design aim to improve the nucleus domains activity and facility in the zinc finger nucleus. Directed evolution produced a variant called as the SHARKEY while structure based design enhances cleavage specifically by modifying the dimerization interface.

So, what happen is that in a in a in a typical zinc finger nucleus what will happen is that it is actually going to utilize the 3 or 4 the zinc finger domain and they are actually going to be utilized a particular stretch of the nucleotide. And once that happens then the zinc finger cleavage domain will actually go and bind to the opposite side of the DNA and it is actually going to cleave and that is how you are going to have the two part right. And then these parts are actually going to be utilized for and will will participate into DNA repair and ultimately it is actually going to lead to the genome editing. Now, what is the application of the zinc finger nucleus? So, zinc finger nucleus as I said you know it is going to be utilized for the genome editing. So, you are what you are going to do is you are going to have the zinc finger nucleus you are also going to have a DNA cleavage domain and then you are going to utilize that for bringing the cleavage into the genome right.

And then this will go for the genome genome repairing right. So, it is actually going to activate the repair mechanisms. So, repairing could be by the non homologous recombinations or by the non homologous recombinations and once that happens it is actually going to bring the changes into the corresponding or the resulting genome. Now, let us move on to the next tool and next tool is the TALEN, Transcriptional Activator Like Effector Nucleus. So, TALEN and zinc finger are actually going to have the approach wise they are same actually.

In the in this case you are using the zinc finger domain as a weapon to identify the DNA where you are actually going to use the transcriptional factors. So, the TALEN as the name suggests transcription activator like effector plus nucleus right. So, transcription activator like effector nucleus are type of engineered restriction enzyme designed to target the specific DNA sequence for a cleavage. They are created by combining a TAL effector DNA binding domain with a DNA cleavage domain which act as a nucleus capable of cutting the DNA strands. By manipulating the transcription activator like effectors, scientists can

customize their binding to the virtually any desired DNA sequence.

When these engineered TALEs are coupled with nucleus precise DNA cleavage occur at a predetermined locations. This engineered restriction enzyme can be introduced into the cell for applications such as genome editing or in-situ genome editing or the method known as the genome editing with a engineered nucleus. TALEN alongside zinc finger nucleus and KSPER9 stand out is the prominent tools in the field of genome editing. So, TAL or the Transcription Activator Like Effectors or TALEs are the protein released by the certain beta and gamma proteobacteria which being a prominent group particularly plant pathogenic xanthomonas bacteria utilize the TALEs produced through the type 3 secretion system. These proteins have a capability to attack to the promoter sequences within the host protein stimulating the expression of plant gene that facilitate the bacterial infection.

Transcription Activator Like Effector protein contains a nuclear localization sequence or NLS, N-terminal localization sequence or the TS, Transcription Activator Domain AD and the central repeat domain or the repeat variant domains. So, the primary distinguishing feature of TAL effector is a central repeat domain consist of 1.5 to 3.5 repeat each typically contain the 34 received long. So, you have a N-terminal residue and then you are owing to have the nuclear localization sequences the N-terminal local translocation signals transcriptional activator domain and central repeat domain.

And in the central you are going to have a central repeat domain and this central repeat domain is unique to the TALEs. The standard repeat sequence is this sequence and which hyper variably residue at the 12 and the 13 position known as the repeat variable dye residue or the RVD. So, this is the region this is the place where it is actually going to show you the variation. The identity of the two residue in consistently repeat correspond to the sequential DNA bases in the TALEs effector target site. According to the crystal structure each repeat comprises two alpha helices and two short RVD containing loop.

In this loop the second residue of the RVD establishes the contact while the first residue stabilizes the RVD containing loop. TALE effector target site commonly include a thymine next to the 5 prime base target by the first repeat. Attributed to the contact between the thymine and the constructed tryptophan in the region N-terminals of the central repeat domain. So, this central repeat domain is going to have this particular sequence and within this the nucleotide number 13 or the residue number 13 and 12 and 13 is very very important because it is actually going to repeat and it is going to be a part of the RVD and that is actually going to decide the sequence identity or the sequence specificity. The experimentally validated code between the RVD sequence and the target DNA bases can be expressed as follows.

So, if you have the A, C or all those kind of thing then it is actually going to have the 5 methyl cytosine and so on. So, on the other hand it is also going to have the C-terminal domain which is a functional domain and is going to act as a activator or repressor or

nucleases or methylases and integrases. So, TALEN will have the 2 domains like first is the TAL DNA binding domain and the second is the DNA cleavage domain. This is exactly the same as what we have discussed for the zinc finger nucleases right. So, TAL effectors the DNA binding domain comprises a repeated 33 to 34 amino acid sequence with variable 12 and 13 base pair amino acid known as the repeat variable digestive or RVD.

This RVD variation correlates strongly with the specific nucleotide recognitions allowing the engineering of the customized DNA binding domain by the selecting the repeat segment with the desired RVDs. Minor alteration in the RVD and the incorporation of non-conventional RVD sequence can enhance the target. So, you can actually be able to alter the DNA of the TAL simply by modifying the RVD sequences or simply incorporating the particular type of RVD sequence. Then we have the non-specific DNA cleavage domain from the foci endonuclease is utilized to construct the hybrid nucleases demonstrated activity in each plant and animal cell. The foci domain operate as a dimer requiring the two construct with the distinct DNA binding domain for the target genome site with appropriate orientation and spacing.

The distance between the tail DNA binding domain and the foci cleavage domain as well as the spacing between the individual TAL domain binding site a crucial factor influencing the activity level. So, this is what exactly happened you have a TAL which is actually going to have the effector function and you are also going to have foci nucleases. So, what you have done is you have fused these two and after the fusion it is actually going to be TAL nucleus right which is actually going to have the DNA binding domain it is going to have the DNA cleavage domain. Now, what will happen is that this portion is actually going to utilize for recognizing a particular DNA sequence right. So, it is not only the DNA, but it also going to recognize a particular DNA sequence and how that happens that happens because of the RVD region where you are actually going to.

So, if you are changing the RVD region within this the central domain it is actually going to recognize a particular DNA sequence and then you also have the DNA cleavage domain. So, it is actually going to cut the DNA and then once it is been cut it is going to go through with the homologous or non homologous recombinations and ultimately it is actually going to give you the altered genome. Now, let us talk about the applications. So, applications of the TALENs are going to be the same as what we have discussed for the zinc finger nucleus and they have very very diversified applications in the various fields of the biotechnology. Then we have the third target or the third tool that is the CRISPR-Cas and CRISPR-Cas is the one of the most versatile and most economically very feasible tool which people can use.

So, CRISPR-Cas is a kind of a immune response. So, CRISPR-Cas is called as clustered regulatory interspaced short palindromic repeats and CRISPR-Cas systems are sophisticated adopted immune system employed by the prokaryotes. CRISPR is a DNA sequence found in the prokaryote like bacteria and archaea originating from the previous bacteriophage infection and served as a adoptive immune system allowing the cell to detect

and element the similar threat. It is present in about the 50 percent of the bacterial genome and nearly 90 percent of the archaea genome. CRISPR consists of the conserved repeat sequences interspersed with the short spacer sequence. CRISPR associated protein or Cas is a enzyme that uses the CRISPR sequences as a guide to recognize and open up of the specific strand of the DNA that are complementary to the CRISPR sequences.

So, CRISPR is a kind of a immune response which bacteria is using to fight against the viral infections. And it is been discovered by the two scientists which are actually been offered the Nobel Prize into the year of 2022, 2020. So, these two scientists have been offered the Nobel Prize in the year of 2020. And there is a long list of the history how the CRISPR-Cas system was evolved. So, it started from in the first time as a 1987 the people have when the people have observed this particular phenomena into the E.

coli. And then ultimately it is been end up in 2024, 2014 when the first CRISPR patent was granted to the feng jiang. So, that is how it is a complete systematic way in which the scientists have discovered the different aspects of the CRISPR-Cas and that is how the complete tool was evolved and this tool can be used for genome editing. So, CRISPR features the short conserved repeat sequences intersect with the similarly sized spacers ranging from the 23 to 47 base pair for repeat and 21 to 72 base pair for spacer. Bacterial genome features multiple CRISPR loci each with the diverse and hyper variable spacer sequence even among the closely related strain. This unique spacer sequence originates from the viral or plasmid DNA and serve as a recognition element.

The introduction of the new spacer allow the system to recognize and eliminate the matching viral or plasmid genome. The CRISPR loci include a conserved leader sequence and their activity depend on the adjacent CRISPR associated Cas genes which encode the essential Cas like protein like the nucleases helicases and polymerases working collaboratively for a CRISPR immune system. So, this is what is the combination you have the CRISPR gene Cas here you have a leader and promoters and then these are the CRISPR repeat spacer array and all these are being utilized for identifying the genome or the DNA sequence from the virus. So, that in the general mechanism you have the 3 events one is the adaptations you have the expression and the maturations and the third is the interference or the targeting. So, in the first event that is the adaptation the CRISPR Cas is actually in the adaptation stage the Cas 1 and Cas 2 complex acquires a proto spacer from the invading viral DNA and integrate it as a new spacer into the CRISPR array.

So, this is what exactly happened that when the virus is going to infect the bacterial cell it is actually going to inject its genome and that genome is going to be identified by the CRISPR Cas 1 and Cas 2 system and it is actually going to generate the proto spacer and that proto spacer is going to integrate into the into the CRISPR array and once that happens it is actually going to you know enter into the next phase and that next phase is the expression and the maturation phase. So, once it enter into this in the expression and maturation phase the CRISPR array undergoes the transcription and it is processed into the

mature CRISPR RNA each contains the transcription spacer and the part of basically repeat these CRISPR RNA form the nucleoprotein RNP complexes with the Cas proteins. And then we have the third phase and third phase is called an interfero phase. So, in the interfero phase the CRISPR RNA Cas RNP complex identify the invading target DNA through the complementary base pair cleaving the target sequence leading to a destruction of the invading viral genome providing the protection against the viral infection into the bacteria.

So, there are different types of CRISPR Cas system. So, you have the class 1 CRISPR Cas you have the class 2 CRISPR Cas and in the within the class 1 you have the type 1 type 2 type 3 and type 4 whereas, in the Cas 2 you are actually going to have the type 2 type 5 and type 6. And all these classification is very very complex right. And the class 1 system utilizes the multiple Cas protein while the class 2 protein involved a single protein for the interference. So, in the class 1 CRISPR Cas you are actually going to have the you know the you are going to have the multiple Cas proteins right. So, you are going to have the Cas 3 you are going to have the Cas 10 you are going to have the other kinds of CRISPR Cas proteins and these class 1 protein and class 1 CRISPR Cas system is going to be further subdivided into like type 1 type 3 and type 4 and type 12 subtypes.

And these class 1 subtype system is found in the 90 percent of the CRISPR loci in bacteria and archaea and can target both the DNA and RNA. So, this is the type 1 where you are going to have the multiple Cas proteins like the Cas 6 and all that and it is actually going to you know target the CRISPR RNA and then it is actually going to enter into the targeting phase where the Cas 3 is actually going to target the particular DNA and it is going to induce the target DNA degradations. So, in a type 1 CRISPR the cascade complex and the Cas 3 nucleus play key roles. The process involves the cleavage of the pre CRISPR RNA by the Cas 6 generating the CRISPR RNA. This CRISPR RNA associated with the cascade identify the proto spacer in a target DNA.

Cas 8 a cascade subunit recognizes the proto spacer adjacent motif. CRISPR PAM function is crucial for a type 1 cascade Cas immune sequence. This functional PAM hinders the CRISPR RNA recognitions impeding the R loop formations. Functional CRISPR system activate the Cas 3 leading to its nick in the target single standard RNA single standard DNA and subsequent degradations. And then we have the type 3 in the type 3 the Cas 6 uses endonuclease endo ribonuclease mechanism to create the CRISPR RNA from the pre CRISPR RNA by cleaving it.

Unlike earlier CRISPR model this type introduces 8 nucleotide repeat sequence called the CRISPR RNA tags through this Cas 6 mediated cleavage. Even downstream to the spacer sequence the CRISPR tag into the 6 nucleotide increasing the size of the CRISPR complex. This is dealt in the Cas 10, CRISPR formation in 2, Cas 3 and Casper complex to B. Type 3 targets both DNA and RNA leading to the coarse transcriptional CRISPR RNA guided cleavage of the target DNA. The PAM domain of the CRISPR Cas 10 cleaves the DNA strands while CSM 3 and the CSR CMR 4 cleaves the RNA transcripts.

And important distinction is that the PAM is not necessarily essential for the type 3 to initiate the immune mechanisms. Then we have the class 2 CRISPR Cas. So, in a class 2 CRISPR Cas characterized by the presence of a single effector molecule. There are 3 types of class 2 systems type 2, type 5 and type 6 and 9 subtypes. While the class 2 systems are more commonly known as Cas 9 is a class 2 system.

They only represent the 10 percent of the CRISPR loci and unlike class 1 they are found in the bacteria. Class 2 system can target both DNA and RNA depending upon the type. Type 2 the most common class 2 system find in type 2. The type 2 systems are characterized by the presence of Cas 9 as well as the ancillary protein like Cas 1 and Cas 2. Cas 9 the commonly used genome engineering endonuclease is a type 2 system.

Type 2 system requires the track RNA for the function. Type 5 system commonly know use the Cas 12 as their endonuclease of choice. Like Cas 9 Cas 12 target the DNA for the editing. Similar to type 2 system the type 5 also require the track RNA for the function. Type 6 is only the last 2 system that target the RNA for editing. So, class 13 is a type of is a type 6 endonuclease that enable the editing of the RNA.

So, then in a type 2 you are going to have the most prominent Cas protein that is a Cas 9 which is actually going to be a part of the type 2 Casper Cas system. And then these are the this is just a summary of the class 1 and class 2. So, in a type class 1 you have a type 1 type 3 and type 5 type 4. Whereas, in the class 2 you have the type 2 type 5 and type 6. In the class 1 you are going to have the signature Cas protein like the Cas 3 Cas 10 and CSF.

Whereas, in the type 2 class 2 you are going to have the class 9 class Cas 12 and Cas 13. So, Cas 9 is very common for genome editing and the Cas 13 is for the RNA. So, these 2 are being used very extensively for editing the genome of the organisms where or the DNA of the organisms whereas, this is being used for the editing of the RNA. So, in a CRISPR Cas 9 system, the CRISPR Cas 9 system particularly the type 2 A is widely studied Cas 9 or 160-kyridol-dal-10 DNA endonuclease is the sole protein needed for the interference. It works in a tandem in the single guide RNA, a fusion of CRISPR RNA and the TRAC RNA alongside the Cas 1, Cas 2 and CSN and Cas RNA 3 for DNA acquisition and processing of pre CRISPR RNA to the mature single SgRNA.

In 2012 this artificial Cas 9 system programmable for any DNA sequence revolutionize the DNA editing. Further research demonstrate the programming capability of Cas 9 from the S-thermophilous Casper system enhancing its application in the target genome editing. So, Cas 9 is actually going to be utilized very extensively into the CRISPR Cas 9 into the type 2 CRISPR system. Then we have the CRISPR Cas 9 system consist of 2 lobes. So, CRISPR Cas 9 system consist of 2 lobes in the Cas 9 proteins, the recognition and the nucleus lobe.

So, in the recognition lobe end you have the residue 62178, 718 and the nucleus lobe

which is actually going to be from 1 to 60 and 719 to 1368. The recognition lobe contain the Rec I, Rec 2 and the domain responsible for the nucleotide recognition. The arginine rich bridge helix act as a linker between the RuVc 1 and the Rec domain playing a crucial role in initiating the cleavage activity upon binding to the target DNA. The nucleus lobe contain the 2 endonuclear domain the HNH domain and the RuC domain. And the HNH domain cleaves the target DNA strand while the RuC domain cleaves the non target DNA strands.

The PAM interacting domain confers the PAM necessity and initiate binding to the DNA upon DNA binding the positively charged residue in the interface between the Rec and NUC lobes specifically at the bridge helix stabilizes the negatively charged SgRNA DNA hybrid. Additionally, positively charged residue in the linker range region between the RuC and NHC domain contribute to stabilizing the displaced non target DNA. So, this is the structure of the Cas 9 where you have the different types of domain you have HNH domain you have Rec 2, Rec 1, RuC 1 and PAM interacting domain and all of these are participating into the CRISPR Cas 9 mediated genome editing. So, what is a guide RNA? So, the these guide RNA take one of the two forms a synthetic trans activating CRISPR RNA or TRAC RNA plus a synthetic CRISPR RNA or CRISPR-CR RNA designed to cleave that gene target site of the interest. A synthetic or the expressed RNA guide DNA sorry RNA a single guide RNA a gRNA that consists of the both the CR RNA and TRAC RNA as a single construct.

So, you have the wild type Cas 9 and you have the modified Cas 9. So, modified Cas 9 is being utilized. So, in the modified Cas 9 you are actually having a D10 to A mutation rendered the RuC domain inactive causing this nickase to exclusively cleave the complementary or target RNA strand. On the other hand the H840A mutation in the HNH domain result in a nickase that cleaves the non complementary or non target DNA. Similarly, you have the dCas9. So, point mutation of D110A and the H840A can deactivate both RuC and HNH nucleus domain leading to a nucleus dead Cas 9 or the dCas9 molecule incapable of cleaving the target DNA.

However, the dCas9 molecule maintain its capacity to bind the target DNA through the guide RNA targeting sequences. So, this is the Casper Cas 9 of the class 12A. So, this is the class 12A system. So, you this is a comparison of the Casper Cas 9 and the Casper CPF 1. So, this is big in size this is complex and Cas 9 required the CR RNA and the TRAC RNA and Cas 9 contain the two nucleus domain HNH domain and RuC nucleus domain.

So, it is 5 prime NGG where N is the any nucleotide and G rich sequences. In this case it is a small in size it is simpler it is CPF 1 require only the CR RNA and CPF 1 contain only one nucleus domain or the RuC domain. And the recognition sequence is also different in this case you have the NGGG where N is any nucleotide whereas, in this case it is TTTN where N is any nucleotide whereas, T is a T rich spam. So, these are the some of the differences between the CR Cas 9 versus CR CPF 1. So, CRISPR is a very very robust tool for genome editing and it has the enormous applications every day there are people who are

utilizing or there are scientists who are utilizing this for genome editing applications. So, it is been employed in you know in genome editing of the plants or virus or the even for the human being also there are reports from the China that it is been utilized for genome editing of the some of the you know new born babies to make them resistance against a particular disease.

So, the CRISPR Cas is a very very robust tool for genome editing and we are going to discuss more about the application of the CRISPR Cas as well as the zinc finger nuclease and the TALENs in our subsequent lectures. So, with this I would like to conclude my lecture here you know subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you. Thank you.

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 12

Applications of Molecular Biology

Lecture-49 Applications of Molecular Biology (Part 1)

Hello everyone, this is Dr. Vishal Trivedi from Department of Bioscience and Bioengineering IIT Guwahati. And what we were discussing, we were discussing about the different aspects within the course called molecular biology. So, so far what we have discussed, we have discussed about the different aspects of cell biology, we have discussed about the central dogma of molecular biology. So, within the central dogma of molecular biology, we discussed in detail about the replications, transcription and translations. And in addition to that, we have also discussed about the different types of molecular techniques which are being developed.

And continuing that discussion, we have also discussed about the genome editing in the previous lecture. So, now since we reached to the end of this particular course, we have decided that we should also discuss about the potentials of studying the molecular biology and how that can be helped to develop the different types of techniques and different types of products which can be very very helpful for the human society. So, if you see the molecular biology, the molecular biology is a very diversified you know field. So, it actually has a different types of tools and techniques and all those kind of avenues and that is why it has extensive applications in the different fields of the biotechnology or in general it is having application in the so much diversified field that it is very difficult to incorporate all the possible fields.

But what I have listed here is only the fields which are very very common and very very popular. So, for example, the molecular biology has a very extensive two application in the agricultural field, pisciculture, poultry, vaccines, it is been used extensively for developing the transgenic animals, different types of medicines, then it also been used for developing the genetically modified organisms and drug delivery. So, since the applications are so much diversified it is very difficult to incorporate or and discuss all these applications in a couple of lectures. So, I have decided that I will going to only focus on the three or four different aspects. So, that it will give you an idea that what will be the potential of the molecular biology and what could be done actually with the using the this particular technique.

And then with this brief discussion it could be possible that you may be able to understand

the importance of the molecular biology for the human society and it also can help you in designing and developing your own product. So, what we are going to focus we are going to focus on to the four aspects. We are going to focus on the genetic engineering, we are going to focus on the PCR based applications, we are going to focus on the how the molecular biology is been used for developing the different types of transgenic animals and how these transgenic animals are having the role in the different aspects related to the whether it is related to poultry or whether it is related to other fields. And then we also going to discuss very briefly about how the genome editing approaches can be having the applications in the diversified field and how that can be used. So, let us first start with the genetic engineering and then we are going to discuss about the transgenic animals and then we also going to discuss about the PCR based applications and then we also going to discuss about genome editing.

So, as I said in the genetic engineering this we have discussed in detail that what is the genetic engineering. The genetic engineering is that you are actually going to engineer the DNA and you are actually going to allow the development of the genetically modified organism. And this is actually the general schemes which actually going to be follow where you are actually going to isolate the particular gene from the genome. And if you recall in previous module we have discussed about how you can be able to isolate a particular gene fragment either utilizing the genomic library or the cDNA library or utilizing the polymerase chain reactions. And then once you got the gene then you are actually going to digest that with the restriction enzyme that is actually going to generate the cohesivants.

And then once the cohesivants are being generated the similar procedure you have to follow for the vector also. And then you are actually going to get the cohesivants both on the vector and as well as on your insert then you are going to put them into the ligation reactions. And the post ligation you are going to transform that or you are going to deliver this DNA into the host cells. And then you are going to screen and select the desirable clones and then ultimately these clones are ready for the future development or future applications. So, ah as far as the genetic engineering is concerned these are the tools which are required for doing the genetic engineering and these are the important procedures which are actually been important for doing the genetic engineering.

Where you are actually going to do the isolation of genomic DNA, polymerase chain reactions, restriction enzymes, ligations and all that. And once you are done with this you are actually at the end of this you are actually going to have the products right. You are actually going to develop the products or you are actually going to generate the enzymes which are actually going to have the diversified applications in the different industries. So, ah if you talk about the dairy industry right ah the dairy industry is extensively been using these products which are been developed by the genetically modified organisms. So, they are actually producing the different types of proteins or they are also producing the different types of enzymes and these enzymes are been having the extensive applications in the food industries.

So, you can have the enzymes which are having the extensive role in the dairy industry, you are also having the enzyme which are having the role in the brewing industry, baking industries, wine industry and as well as the meat industries. And we will not going to discuss in detail about these enzymes and how they actually going to perform the different types of functions and how they are actually going to be useful, but ah very briefly we will talk about what are the how these enzymes are ah having the application in the particular field and how you are actually going to be produce those enzymes. So, first we start with the dairy industry. So, in the dairy industry we have the 4 enzymes which are called rennet, lactase, protease and catalyst and all these are actually having the different types of applications and all these enzymes are genetically been cloned into a particular over expressing ah cell and that is how they are actually been produced on to the industrial scale and then they are actually been used into the dairy industries. So, as far as the rennet is concerned ah it is extracted from the stomach of the young calf and apart from that it also been developed by the ah recombinant DNA technology.

This contains the enzyme that cause the milk to become the cheese. So, it is actually an enzyme which converts the milk into cheese and it separates the solid curd and the liquid whey and the different animal rennets are also been used for the different cheese and most common vegetable rennet is thistle ok. So, rennet is an enzyme which is actually been used for converting the milk into the cheese and that is how they are actually going to help in ah in the in the dairy industries. Then we have the lactase. So, lactase is present in the brush border of the small intestine it is artificially extracted from the yeast and it is required for the digestion of the whole milk and it is used in the production of lactose free ah milk right.

So, you know that the milk is ah is a dairy product which contains the sugar right and the sugar component of the milk is always the lactose and you know that the lactose intolerance is a very very big issue because if you if the somebody is lactase insensitive ah then it is actually going to develop ah it this lactase will lactose is not going to be digested. Then as a result the lactose is going to be remain undigested and remain in the stomach and that is how it is actually going to cause the production of gas and bloating and all those kind of things. So, to avoid that you are actually can get the milk which is free of lactose and how you are going to do that you are actually going to ah digest the milk with the help of an enzyme which is called as lactase. So, what the lactase is going to do is it is actually going to ah you know take the lactose from the milk and it is actually going to chew up right. So, it is actually going to eat up all these lactose and it that is how it is actually going to make the milk which is free of lactose and that lactose free milk is actually having a very high economical values compared to the normal milk because that milk can be given to the patients, it can be given to a special people who are actually having the lactose sensitivity and lactose intolerance and all that.

Ah It is also been used in the production of ice cream and the sweetened flavor and the condensed milk. Then we have the catalyst. So, it is catalyst is produced from the bovine

liver and the microbial sources. It break down the hydrogen peroxide to the water and the molecular oxygen and along with the glucose oxidase it is used in the treating the food wrappers to prevent the oxidations and it is also been used to remove the traces of hydrogen peroxide in the process of cold sterilizations. So, the catalyst is very very important for detoxifying the hydrogen peroxide which is going to be residual be present into the food products because many of the food products are been sterilized with the help of the hydrogen peroxide and the hydrogen peroxide is toxic.

So, to remove the hydrogen peroxide you are just adding the catalyst and the catalyst what catalyst is going to do is it is going to convert the hydrogen peroxide into the water an oxygen. Then we have the proteases. So, proteases are the general enzyme they are actually been used in the many industries they are been used in the you know the dairy industries, the meat industries and all that. So, proteases are been widely been distributed in the biological world and they hydrolyzes the specific peptide bond to generate the paracase paracasein and the macro peptide in the production of cheese and that results in the bitter flavor to the cheese and also in a desired textures. Then let us come to the brewing industries.

So, within the brewing industry as I said you know protease is very very common with within the dairy within the different types of industrial setup. Then we can have the beta glucanase, we can have the alpha amylase and the amylo glycosidase. So, protease is protease works to provide the what with the amino acid nutrient that will be used by the yeast right. So, and protease work to break up the large proteins which enhance the head retention of the beer and reduce the haze and it fully modified the malt these enzyme have done with their work during the malting process. Then we have the beta glucanase.

So, beta glucanase represent a group of carbohydrate enzyme which break down the glycosidic bond within the beta glucan. It aids in the filtration after the mastication and the brewing. Then we have the alpha amylase. So, alpha amylase convert the starch to the aedestrine in producing the corn syrup and it stop solubilizes the carbohydrate found in the barley and other the several cues in the brewing. It decrease the time required for the mashing.

All these technical terms like mashing and brewing and all those we are not going to discuss and neither I expect that you should know all this, but if you are interested you can read about these content and read material in somewhere else. Then we have the baking industry. So, in the baking industry you can have the maltogenic amylase, you can have the glucose oxidase and you also have the pentonases and all these are actually having the diversified function and the applications within the baking industries. So, maltogenic amylase. So, it is a floor supplements, it has anti-stalling effect, it modify starch while most of the starch starts to gelatinize.

Resulting starch granules become more flexible during the storage. Then we have the

glucose oxidase. Glucose oxidase oxidizes the glucose and produce the gluconic acid and the hydrogen peroxide. hydrogen peroxide is a strong oxidizing agent that strengthen the disulfide and non-dulciferide cross link in the glutens and it good working condition help proper functioning of the baking bakery systems. Then we also have the pentonases.

So, its exact mechanism is not yet discovered. It improves the dough mechanability yielding a more flexible easier to handle dough. The dough is more stable and give better oven spring during the baking. Then we also come to the wine industry. So, in the wine industry we have the two enzymes called the pectinases and the beta gluconases which have the extensive role and both of these enzymes can be produced with the help of the recombinant DNA technology.

So, pectinases prevents the pectin from forming the haze and hence to get the clear solutions. Similarly used for the extraction of color and juice from the fresh fruits, it breaks down the pectin and releases the methanol and high amount is hazardous. Then we also have the beta gluconases and it accelerates all biological mechanism linked to the maturation of leaves. It reduces the maturation durations and it improves the clarification and filtration and improves the action of the active finding regions. Then we come to the meat industry.

So, within the meat industry we have the two enzymes which is called as the proteases the papain. The proteases or the papain both are actually being used extensively in the meat tenderizations ok. So, meat tenderization actually enhances the economic value of the meat and also enhances the taste of the meat and that is how it has been very very desirable and both of these enzymes are being cloned and over expressed and that is how they are being used in the industry. So, first is the proteases. So, proteases break down protein which will disrupt or loosen the muscle fiber and tenderizes.

So, when you treat the meat with a protease it is actually going to you know break the peptide bonds and that is how it is actually going to make the meat little soft and easy to digest and it also gives some different flavor and that is how it is actually going to be used in enhancing the taste of the meat and that is how it has been desirable to have this particular type of enzyme in large quantity and that is how they are being used ah they are being produced with the help of the recombinant DNA technology. Then we have the papain. So, papain which is found in papaya 95 percent of the meat tenderization available in grocery stores are made from the papain it is extracted from the latex in the papaya fruits and these enzymes are purified and sold in the powder or the liquid form and that is how they are actually been used ah. Now, let us come to the another field where the product of the genetic engineering can be used. So, another industry is the medicinal ah medicinal world ok.

So, in the medical medical world the ah the product of the recombinant technology is been

used extensively as a drug, as a vaccine, as adjuvants, as in the gene therapy and all those kind of things and all these are actually requiring the knowledge of the molecular biology. So, let us first discuss about the applications in the medicinal science. So, medicines are class of molecule used to correct the disturbance in the host physiology. They can be chemically in nature and used to inhibit the abrid enzyme activity from the host or pathogen. In few cases the host enzyme can be supplied as a drug formulation to drive the biological reaction.

Biotechnology has a potential to contributing into the development of the drug molecule and biotechnology means the genetic engineering right. So, with the help of the genetic engineering the combinatorial technology you can be able to clone these ah proteinaceous substances or the enzyme and then you can be able to supply. So, you can actually have the four different classes. One is the production of therapeutically important proteins, you can actually be able to do a gene therapy, you can be able to develop the vaccines and you can also be able to produce the monoclonal antibodies. And all these aspects require the extensive knowledge of molecular biology so that you can be able to manipulate the cells, you can be able to clone a particular gene and you can be able to use that for the different types of applications.

So, let us first discuss about the production of the typically important proteins. So, a large number of genetic or metabolic diseases can be corrected by supplying the protein or the factors. Following the advancement in the biotechnology many other proteins or factors are produced in the different bacterial expression system. In an approach the gene of the enzyme or the proteinaceous factor is cloned into the appropriate plasmid to produce the recombinant clone. One of the such example is the human insulin right.

Human insulin is one of the widely ah you know supplied ah biological product right and you know that the diabetes is a very big disease right and the human insulin is a is actually been required to lower down the blood glucose and that is how it has a very huge ah market in which you can be able to use. So, earlier when the recombinant technology was not known, the human insulin is always been isolated from the animal sources and then they are actually been given to the patient and the major drawback of this particular approach is that since you are giving the insulin from the ah from the animals, they may actually have the allergic reactions or they may actually not be get accepted by the human system right and that is how in those cases in those era the insulin was creating lot of problems. So, to ah correct that particular problem what people have done is they have a when the people were you know ah recombinant known the recombinant technology how they will know how to clone the ah gene and how you can be able to use that for producing the protein, what they did is they have cloned the insulin. So, insulin is a protein of two ah chain, you have a chain, you have a b chain and both of these a and b chains are connected with the help of a disulfide leakages right. So, what you can do is you can produce the a chain, you can produce the b chain separately and then you just combine them and that is how it is you are going to get a functionally active the insulin molecule.

So, what you are going to do is you are going to accelerate the gene a, you are going to accelerate the gene b, you are going to transform that into the bacteria and that is how you are going to get the gene ah peptide a right and you are going to get the peptide b and then ultimately you are going to mix them, you are going to do the you know the you are going to change the conditions in such a way that the disulfide linkages are going to be formed and that is how you are actually going to have the functionally active insulin. So, that is what it is written here that insulin is a dimer of a chain a and b chain linked by the disulfide bonds composed of the 51 amino acid with a molecular weight of 5808. A schematic representation step of given right in this particular figure in this process the gene a and b is cloned into the bacterial plasmid separately to produce the two recombinant clones. Peptide chain a and b is over expressed in the E. coli and recombined together to produce the functional insulin.

So, this is just a simple example to show you the potential of molecular biology potentials of the combinatorial technology to say that how something which was very very crucial for the human welfare was being done with the help of the molecular biology. Earlier people were you know getting the lot of side effects, side reactions and all those kind of things. So, apart from insulin there are so many different types of hermetically important molecules are being generated with the help of the molecular biology. So, what we have is we have the protein ah factor 8 right and the factor 9 these are the factor which are required for the black clotting. So, they are actually been required for the treatment of the hemophilia.

Then we have the tissue plasmagen activator that is also been required for the thrombosis then we have the lactoferrin. Lactoferrin is been used for treating the GI tract infections then we have the human protein C that is required for thrombosis. Then we have the alpha 1 antitrypsin that is for the inphysema then we have the fibrogen. Fibrogen is required for the wound healing right. So, in some cases what happen is that when you get the wound you are not producing the enough quantity of fibrogen right.

And you know that the fibrogen is been get converted into the fibrin and these fibrin fibers are actually covering the wound and then only it is actually going to work to you know to heal the wounds. So, in those cases you are actually going to do the you know the you are going to use a recombinant tetralogy to produce the fibrogen and that fibrogen can be used for wound healing. Then you also have the pro 540 5 4 5 4 2 and that is been used for the HIV infections. Then we have the anti anti thrombin 3 that is for correcting the thrombosis. Then we have the collagen 1 that is been required also for the tissue repair and then we also require the serum albumin that is for increasing the blood volume ok.

This is very important right because if you are running with the low in protein you your blood volume is also going to be reduced to maintain the tonicity and as well as the osmobilartie. So, if you increase the protein volume into the your blood the blood volume is also eventually going to increase. So, apart from that you can also have the recombinant

chymosin and the recombinant human growth factors and recombinant blood clotting factor 8. So, these are some of the proteins which are required for the different types of applications. Now apart from this you can also have the you can also be able to do the gene therapy with the help of the molecular biology.

So, production and supply of recombinant protein is a temporary solution for a treatment of a disease condition. In another approach the human expression system is used to produce the proteinaceous factor after inserting the recombinant clone into the human cell or inside the human body. And DNA is packed into the appropriate DNA delivery system like you can use the viral system, you can use the liposomes, you can make all those kind of things. To deliver the gene into the human cell to correct a mutated gene or encode a therapeutic protein drug to provide the treatment. So, gene therapy is also very very common and very very popular method through which you can be able to correct the problem at the molecular level right.

Now what you have seen that when we are genetically you know producing the proteins you are actually supplying only the protein part, but that protein is having the half life right. So, that protein gets disappear after some time, but in this case what we are doing is we are putting the DNA. So, once we are putting the DNA into the cell, DNA will keep producing the protein for a long longer period of time. And if it is a permanent transfection, if it is a permanent integration then it actually going to also change the cells and then the cure is going to be permanent. So, there are two different types of gene therapy you can have the somatic gene therapy or you can have the germline gene therapy.

So, somatic gene therapy is for those cells which are like somatic cells like for example, muscles, liver, pancreas or those kind of thing. And then you can have germline gene therapy in this the therapeutic approach the germline cells like the sperm or egg are transformed by the introduction of required gene to produce the protein or correct the mutated gene. So, this is actually going to be done in the sperm as well as the egg. So, that the offspring the corrections are not going to be done in the that current gene current generation, but it also going to be done into the future generations. And in that case if this particular correction is going to be done for the several generations.

Whereas, in the case of somatic it is only going to be for that individual where you are doing the gene therapy. Whereas, for the germline gene therapy it is also for the incoming generation as well. So, the technical problem what is associated with the gene therapy is that it is short lived because as I said you know you are going to insert the DNA into the cell and recombinant DNA may express may not express or it may actually get rejected by the system. Then we also have the immune reactions because you are injecting the virus containing gene and all that.

So, it is also can have the viral infections viral reactions. Then since you are using the viral vectors which can cause the you know the immune reactions adverse immune reactions

and toxicity. And then it also can disturb the human physiology because it also the gene in to get into a wrong place in the genome it may cause the functional defects. And this is what we have discussed when we were discussing about the homologous and non homologous recombination as an approach to for the genome editing right. So, this is what one of the approach where you are actually going to put the flanking sequences as a homologous recombination homologous sequences. And then utilizing these flanking sequences the gene of your gene what is present on to your vector is going to be inserted into the vector into the genome of that particular cell.

But if any of these events are go wrong then the gene may get integrated into a off target sites and that actually is going to cause the problem into the host physiology. Because then it is actually this see it is not going to correct the problem, but it is also going to make the additional problem because the it may replace some of the gene which was working correctly. Now, apart from this you can also be able to use the recombinant technology for producing the vaccines. So, vaccine is been given to develop the immunity against a disease in a human or the vertebrate animals. Vaccines are of different types like the dead or attenuated organisms or the protein derived from them.

There are different strategies to enhance the immunological responses to give the long lasting protection against a disease with the minimum adverse effects. There are four different types of vaccines you can have the killed vaccine, you can have attenuated vaccine, you can have a toxoid, you can have a subunit vaccine and you can also have the conjugate. So, all of these approaches are requiring the one or other tools of molecular biology for developing the vaccine. You might have seen the COVID vaccine how the people have utilized the molecular biology tools to develop the COVID vaccine in a small very short span of time and that is the potential of the molecular biology. So, this is what we have discussed in the in the genetic engineering.

So, within the genetic engineering you are actually going to produce the proteins or you are actually going to produce a recombinant DNA and that can be used for generating the different types of plasmids or different types of products. And that different types of product can be used for treating the different types of diseases or it can be used for correcting the particular type of errors into the metabolic reactions. So, with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss some more applications of the molecular biology. Thank you.

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 12

Applications of Molecular Biology

Lecture-50 Applications of Molecular Biology (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioinformatics IIT Guwahati and what we were discussing? We were discussing about the different aspects of the molecular biology in the course molecular biology. So, far what we have discussed? We have discussed about the cell biology, we have discussed about the molecular biology, we have discussed about the central dogma of molecular biology, we have discussed about the transcription, translation, replications, post translational modifications, we have discussed about the cell cycle, cell division, apoptosis and all other kind of aspects. And apart from that we have also discussed about the different types of techniques what has been developed. So, we discussed about the polymerase chain reactions, real time PCR, blotting techniques, we have discussed about the southern blotting, northern blotting and western blotting.

And in the previous module we have also discussed about the genome editing. So, and we have discussed about the different aspects of the genome editing in very briefly we have discussed all these. So, with this we have discussed in detail about all the tools and techniques which are available with the molecular biology. And now it is a time to review to see how these techniques can be utilized for the welfare of the human society.

So, what we have what we have discussed we have discussed about that molecular biology can be can have a extreme potential into a different heels. It can be used for the poultry industry, it can be used for the agriculture, it can be used for generating the transgenic animals, it can be used for developing the recombinant proteins and so on. So, since the scope is diversified we decided that we will focus only on to the four aspects related to the molecular biology. We are going to discuss about the genetic engineering, we are going to discuss about the PCR based applications, we are going to discuss about the transgenic animals and we also going to discuss at the end about the genome editing. And how the genome editing can have the applications in the different types of areas.

Now, in today's lecture we are going to discuss about the transgenic animals. So, transgenic animals so, when we say transgenic animals these are actually the animals where you are

actually going to have the two different types of characters two different types of characters what are present. And transgenic animals are the animals where you are actually going to put the external DNA into the into the particular right. So, transgenic animal means you are actually going to put the genetic content from the trans which means from the outside and that is how it is actually going to modify the particular animal. So, what is the application of the transgenic animals? The application of transgenic animal is that it is actually going to be used for understanding a particular disease mechanism.

So, we will take a few examples then you understand how the transgenic animals are utilizing for developing the different types of disease mechanisms. One of the classical example is the cancer for example. So, if a particular cancer is been developed because the particular type of transcription factor is been required and it is actually over expressing in a large quantity. So, then what you can do is you can just over express that particular transcription factor into a particular animal and that animal is eventually going to develop the cancers of the breast or pituitary or any kind of thing. So, that can be used for studying the how the cancer is been developed and that can also be used for the studying the or exploring the anti cancer drugs also.

Then it also can be used for studying the normal physiology. So, just like as we discussed for the cancer it can also be used for studying how the blood clotting is working, how the nerve conduction is working, how the all the muscle contraction all those kind of. So, because all these mechanisms are requiring the different types of factors. So, if you remove one or two other factors it can help you to understand what will be the role at that particular factor. For example, if you remove the actin and if you remove the myosin or if you remove the some crucial GTPases it will tell you that ok, this is the function of the particular protein and that is how this is going to help you understanding the normal physiology.

Then it also can help you in producing the biological products. So, biological product where you actually can produce or you can actually be able to generate a transgenic animals and that is good in giving you a particular desirable product. Then it also can be used for vaccine development and it also can be used for the toxicity testing. For example, toxicity testing in those cases where you cannot test the particular product directly onto the human, but you can actually be able to develop the humanized animals, you can be able to develop the transgenic animals which are actually going to mimic the physiology of the humans, but they are not humans and then you can actually be able to do the toxicity testing in on those animals. So, as far as the transgenic animal is concerned transgenic animals could be of two types.

It could be a animal where you are over expressing a particular gene or it actually can be a transgenic animal where you are actually going to remove a particular gene. So, when you are over expressing a gene you are expressing a gene into a suitable vector and these kind of transgenic animals are always been used for the organ transplant. Compared to that

when you are doing the knockouts you are removing a particular gene expressing functional gene and this is useful for understanding the role of that particular gene into a particular process. For example, cancer development, normal physiology and all that. So, first understand the disease mechanism.

So, we will discuss about the transgenic animals and how they are been exploited to exploring the different types of aspects related to molecular biology or other biochemistry fields. So, first is the transgenic animals which are been used for understanding the disease mechanism. So, you can actually have the disease model, you can have the disease model like conventional method where you are actually going to treat a particular animal with a particular type of treatment and that is how it is actually going to have the conventional method. So, in the conventional method what will happen is that you are actually going to treat the animals with like for example, you can actually be able to treat animal with a mutagenic compound. So, if that will happen then it is actually going to cause the mutations or you can actually be able to use the chemicals and that is how it is actually going to cause the development of particular disease.

For example, if you treat the rats with the streptomatose gene it is actually going to develop the diabetes mellitus. Similarly, you can actually be able to use the some kind of physical method. For example, you can actually be able to use radiation and if you radiate a particular organ that organ is actually going to be disappeared and that is how you can be able to develop a disease model. For example, when people have discovered the diabetes right they actually literally remove the pancreas of a particular animal. So, when they when they remove the pancreas from the dogs they found that the dogs are developing the diabetes and that is how these are actually are the conventional method.

What is the limitation of this? That limitation is that there will be no precise control you are actually removing the pancreas. So, if you are removing the pancreas you are not removing only one you are removing a complete organ and that complete organ may have the some more functions right it may just not have the function of reducing the insulin it is also going to have the other kind of effects. Then it also going to cause the gross changes in the physiology right. So, gross changes in the physiology you do not know that if it is directly or indirectly affecting that particular process. Then it may contribution of the multiple organ it could be possible that there could be a contribution of multiple organ in that particular process and you may be just removing one organ or two organs and that is how.

So, to avoid these kind of complications people are developing the transgenic animals. So, that you can very precisely be able to say that ok, the effect of actin in this process is this effect of myosin is in this process is this. So, that is one of the advantage of developing the transgenic animals. So, genetically altered animals where you are either going to over express or knock out to resemble the signature pattern of a human disease conditions. Now, what is the major advantage of developing a transgenic animals? It is target specific

gene site you know that what gene you are trying to you know trying to test right.

So, you can actually be able to do that there is no human subjects involved then it is cheap and quick and it is reproducible. For example, for a disease model like AIDS mouse, alzheimer mouse, onco mouse, diabetes, zebra fish based models and all that they are being developed. Then for example, the AIDS model right. So, no reported case of AIDS in the animal right. So, monkey can get only infected with the SIV.

So, first AIDS mouse model is skid mouse right and the animal used for that model is rat, cat, rabbit and Drosophila. And what is the advantage of animal AIDS model? It is actually cheap, it is much is already known about their genes. So, you know that this animal what is the genetic makeup of this animal and so on. So, it is very difficult it is easy for you to decipher right. And then let us talk about the another model which is called as the Alzheimer's mouse right.

So, it is a neuro generative brain disorder which is where the amyloid plaque or the amyloid beta part peptides fragments of beta are actually going to be accumulated into the brain. And the APP mutation to induce the protein aggregation is the key to develop the AD mouse model right. So, TG2576 is a good transgenic mouse model to start with and in order to produce the genuine human amyloid. Words first Alzheimer mouse incorporated and APP mutations. The pathophysiological effect were similar in the transgenic mouse and it is comparable to the human.

So, Alzheimer disease is a very very serious disease of the central nervous system and it is actually going to be. So, you can actually be able to use these mouse model to screen the compound, you can actually be able to use them for developing the new methods of developing the treatment and so on. And then we have the cancer mouse model. So, cancer mouse models are very very common for example, this is a nude mouse. So, to study the cancer formation and to screen the anti-tumor drugs oncogenes are either mutated to induce the spontaneous cancer formations, knock out of a tumor suppressor gene for example, RB genes or the p53.

Then you can also have the HER2 overexpression and that is also associated with the cancer formation. For example, the HER2 transgenic miles which is called FBB slash N dash MMT new mouse and there are 13 different types of strains which are been engineered to develop the transgenic animals and those in transgenic animals are spontaneously going to develop the cancer. So, they will mimic the natural way of developing the cancer. Then we have a transgenic model for diabetes. So, for studying the gene and their role in peripheral insulin actions, the model of insulin secretion is the glucokinase and the beta receptor knockout mice, uncoupling protein knockout mice and acute and chronic model for the anti-diabetic agents.

Then we also have the zebra fish and the zebra transgenic zebra fish which is for the

cardiovascular disease model. So, these are some of the cardiovascular diseases for which the transgenic zebra fish model is being developed. And what is the major application of the transgenic animal? They are actually being used for understanding the disease mechanism, normal physiology. So, transgenic animals which are genetically engineered to produce a human pharmaceutical product in their saliva, milk, urine or blood. So, they are being called as farming.

Transgenic animals produce the RPs or the products for example, the milk for example, human anti-thrombin 3. Then we also have the chicken egg white. So, for example, the MAB and human interferon gamma NSH etcetera and insulin which is for the diabetes mellitus. And then we also have the production of vaccine into the transgenic animals like the growth factors, coagulation factors and lactoferrin and the infant formula feed. Then we also use the transgenic animals even for the organ transplant right.

So, organ transplant you know that you require a donor organ donor person you actually require a suitable organ. So, that it is matching with the acceptors and so on. So, that is a very complicated process. So, and at the end you require a human being to give you the organs right. So, to avoid that you can actually be able to develop the transgenic animals.

So, that you can develop a particular organ or you can actually be able to take directly from that particular transgenic animals. So, transgenic animals for the organ transplants. So, transgenic animals that are genetically modified to have organ that can be transplanted into the human do not express the key foreign antigens. So, you can actually be able to avoid the graft rejections. Then promising the alteration to the human runners pig is the animal used in this process so far.

Why it is so? Because the pig is big in size thus the size of the organ what you are going to get from the pig is probably will be functionally active. And the pig is actually very close to the human being. So, their physiology and their processes are going to be identical or almost similar to the human. So, that is why it is been accepted or. So, for example, the porcine heart wall successfully transplanted into the human heart.

Then its physiology is closer matching that of human and that is one of the positive point of developing or transgenic animals especially the pig. And it is much less expensive than the monkey and other primates. So, then it is called as super pig and it is fast growing and lean animals it express the bovine growth hormones. So, super pig is a transgenic animal which is been developed to produce the bovine growth hormones. Then we have the super fish which are called as the aqua advantage salmon and these are the genetically modified Atlantic salmon.

Then we also have the GH regulating gene from the Pacific Chincu salmon. Then we also have a smart mouse. So, smart mouse is a superior learning and memory and the doggy

mouse. So, it is modified to over express the NRF NR2B transgene in the cortex and hippocampus and thalamus and brain stem and cerebellum. So, these are the you know the mouse which are actually having the higher IQ values.

Then this is the transgenic animals and their advantages. So, these one of the major advantage is that the gene requires certain cellular mechanism to help the production of protein. The animal used to trans use for transgenic purposes naturally carry the mechanism needed to produce the complex protein. This mechanism is absent in the cell culture. So, when you try to produce the protein in the same culture it happens sometime that the protein what you are producing at the final stage may not be that much efficient compared to that the protein what you are producing in the transgenic animal.

Because in the transgenic animal naturally it is getting into the another the effect from the vicinity and because of that it is may be more effective. The expression to cell culture or bacterial culture require the constant monitoring and sampling right. Because you are supposed to feed the animals your animal cells you are supposed to do all those kind of thing right. Whereas if you are developing a transgenic animals you are actually animal will do all is that right animal is going to supply the nutrient for the liver animal is going to supply the nutrient for all its organ right.

So, that will be easier. The isolation and the purification of expressed protein in a conventional method is more difficult than the purifying protein from the animal or milk from the body fluid. So, that is also one of the thing right. If you are developing a transgenic animals you can actually be able to have the large quantity of that particular protein compared to that if you want to isolate that same protein from a from a non transgenic animals. It is more cost effective as product is efficiently passed through the milk on a average yield of 53 percent and a 99 percent purity. It has been estimated that the transgenic animals can produce it in its lifetime is 100 to 200 million worth of the pharmaceuticals.

But apart from these the transgenic animals advantages you also have the several disadvantages. For example, the transgenic animal project is extremely expensive. Generation of transgenic animal is also expensive because of the long gestation period, litter size and the higher maintenance cost of the recipient animals. There may be a high mortality rate and other deleterious effect on the animal used by the researcher to create the transgenic breeds. It has been observed that the transgenic pigs have enhanced the growth rate and efficient feed convention exhibit reduced reproductive performance and may suffer from the arthritis and the dermatitis.

Large number of recipient is required to amir transfer because of the low transgenic rate and the transgenic food that has been produced and better productivity in terms of both yield and quantity. However, there are some appreciation about the safety of the transgenic foods. So, one of the major challenge is that many people do not accept those product

because they are coming from the transgenic animals right. So, there are ethical concern, there are the social concern. So, that may be are not accepting these transgenic the product from the transgenic animals.

So, these are the some of the advantages and as well as the disadvantages of the transgenic animals. But molecular biology is being used extensively to produce the different types of transgenic animals and they are very very useful in terms of addressing the or developing the disease models. They are being used for developing the understanding the mechanism of a disease and as well as the normal physiology. So, molecular biology has a immense potential to provide the products for the social welfare.

So, with this I would like to conclude my lecture here. In our subsequent lecture we are going to discuss some more aspects of the molecular biology and we will discuss how the molecular biology can be used for developing the different types of techniques based on the PCR techniques and as well as the genome editing techniques. So, with this I would like to conclude my lecture here. Thank you. .

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 12

Applications of Molecular Biology

Lecture-51 Applications of Molecular Biology (Part 3)

Hello everyone, this is Dr. Vishal Trivedi from Department of Biosciences and Bioengineering IIT Guwahati and what we were discussing? We were discussing about the different aspects of the molecular biology in this particular course. So, so far what we have discussed? We have discussed about the basic cell biology, we have discussed about the biochemistry of the different biomolecules and then we also discussed about the central dogma of molecular biology followed by we have discussed about the different types of techniques such as the blotting techniques and we have also discussed about the polymerase chain reactions, real time PCR and in the previous module we have also discussed about the genome editing. Now in current module we are discussing about the application of the molecular biology and if you recall in the previous lecture we have discussed about the how the genetic engineering and the transgenic animals are been you know are been used for the different types of applications and how the molecular biology basic principles are been used to perform the genetic engineering and how the genetic engineering in turn will lead to the production of the different types of products. Whether these products are the enzymes or the different types of proteins which is required for correcting the different types of disorders and along with that we have also discussed about that how the genetic engineering lead to the discovery of the different types of transgenic animals and these transgenic animals are mimicking the disease models they are also been used for the protein production they are also been used for various other applications. So, all these we have discussed in the previous lecture.

In today's lecture we are going to discuss about the PCR based applications and we are also going to discuss about the genome editing. So, let us start our discussion about the PCR based applications. So, in the PCR based applications we are going to discuss about the application of the PCR and as well as the real time PCR or RT-PCR. So, let us start our lecture and discuss about the application of the PCR application of the PCR and as well as the real time PCR.

If you recall when we were discussing about the PCR we have discussed in detail about how you can be able to design the primers, how you can be able to perform the PCR and so on. So, it is better if you can actually be able to briefly go through with those lectures. So, that it will be easier for you to understand the application part. So, in today's lecture we are

going to discuss about the PCR based applications. So, we are going to discuss about the PCR based applications and we are also going to discuss about the RT-PCR based applications.

So, let us start about the PCR based applications. So, when we talk about the PCR based application it is highly recommended that you should go through with the previous lectures. So, that it will be easier for you to follow these content because we are not going to discuss about how you can be able to design the primer, how you can be able to perform the PCRs and so on. So, let us start discussing about first the application of the PCR and then we are also going to discuss about the application of the real time PCR. So, when we talk about the PCR application of the PCR, the application of the PCR lies into the three major components or major areas.

First is the molecular identification, the second is the sequencing and the third is the genetic engineering ok. And when we talk about the molecular identifications that is having the in homeliness applications in the different types of fields. For example, it can be in the molecular archaeology, molecular epidemiology, molecular ecology, DNA fingerprinting, it is also being used for the classification of the different types of organisms. So, it is also being used for the genotyping, prenatal diagnosis, mutation screening, their discovery, genetic matching and as well as the detection of the pathogen. Now when we talk about the sequencing it is also being used for the bioinformatics purposes, genetic cloning and as well as the it is been extensively been used in the human genome project.

And as well as the genetic engineering is concerned it is been used for the site directed mutagenesis and as well as for the gene expression studies. So, let us first start with the first aspect and then we will follow the these are the two aspects. So, PCR is extensively been used in the detection of the different types of reagents or agents in the food science. So, PCR is a rapid and sensitive method that enable the detection of the subdominant population of food without the need of the enrichment media. It allows the detection of the dead cells and the non-cultivate cells which means it is actually going to be used for even assessing the quality of the product.

And then the also been used for detection of the different types of microbes for example, the major infectious organism what is present in the food is the salmolar species. So, in the salmolar when you want to detect the salmolar you can actually be able to detect the INVA enzyme and it can be used for the detection and that can be done with the and it is having a application in the terms of the artificially contaminated the chicken meat or the salmon or raw milk and all that. Similarly, you can have other species of the salmolar which are also going to be detected into the food material and that is having a very huge application because it actually helps in getting the salmolar infection. And salmolar is a very very deadly bacteria. So, it actually can cause the food poisoning it also can lead to the death of the human being.

Then we also have the Listeria monocytogenesis Listeria for detecting the Listeria you are going to use the 16 rRNA sequences and it is going to be used for the detection and the quantifications and it is going to be used for detection of this Listeria in the different types of vegetables like cabbage, lettuce and parsley and the onion branches. Then we also have the Stellophagus aureus that is so, for detection of these Stellophagus you are going to use the NUC gene and it is also been used in the majority of these products are being used for detection of the beef or the other meat products. Then we have the enterobacteria we have E. coli we have bacillus celus groups then we have the you know viable bacteria and norovirus. The norovirus is a virus which actually causes the stomach flu in the kids and for detecting the norovirus you can actually be able to use the ORF1 as the gene sequence and you can be able to detect the contamination of the cheese or the lettuce.

Then we also can detect the hepatitis A virus which is actually the leading cause of the hepatitis in the through the food borne hepatitis and you can actually be able to detect using this with the VP1 to VP3 capsid region that gene and it can be used for detecting of the artificially contaminated the tomato sauce or the blended strawberries. So, PCR has a extensive application in the medical sciences. So, the use of PCR in the medical science already began in the year of 1985 and a viable test for measuring the amount of HIV in the blood was published in the May 1987 by the John at CITAS. In 1989 the team has developed the multiplex PCR and amplification of a single cell was performed for the first time on a single sperm cell to directly analyze the product of meiotic recombination and was also applied to another target cells like the human mucocyte antigen DQ alpha. The PCR technology has become an essential research and diagnostic tool for improving the human health and the quality of light.

It allows the detection of the infectious organism just from one cell by amplifying the specific region of the genetic material. The PCR is very extensively been used for detection of the infectious organisms such as it can be used for detecting of the HIV, hepatitis B and C, human papilloma virus, then we also have the Cladiminus tracheodormis, then the ceria and the cytomegalase virus or the CMB virus and it also can be used for detection of the mycobacterium tuberculosis. Majority of these viruses or the bacterial species are known to cause the different types of disease. For example, the hepatitis B and C is causing the liver cirrhosis and the cancer. Similarly, human papilloma virus which can lead to the cervical cancer, then we have this cladomyelitis which can result in the infertility in the woman, then we have an a ceria which cause the pelvic inflammatory disease, then we have the mycobacterium tuberculosis which causes the tuberculosis.

And then how you are going to do the infection or detection by the PCR, what you are going to do is you are going to draw the blood of the patient and then you are going to perform the PCR with the help of the target gene, the primer designed for the target gene and then you are going to analyze it onto a gross gel and then you are going to make it with the marker. So, it is actually going to say whether if the product is being formed, then you are going to say that this particular sample is actually contaminated with this particular

infectious disease, whereas you are also going to run a control sample where you are going to have the healthy individual. And so this based using this particular type of technique, you can be able to do the blood screening and you can be able to detect the different types of diseases like the hepatitis, B and C and HIV and all that. So, and then we also can use the PCR in the plant science. So, there are various fields in plant science which require the use of the PCR technology for its accomplishment.

One of the major area where the PCR can be used for plants species identifications. So, the PCR technique has also been employed in the identification of the plant species using the species and group species primer targeting the chloroplast DNA. These assay allow the identification of the plant based on the size specific amplicon. For example, plant belonging to the same family has closed primer binding site and hence same amplicon sites where the plant belonging to the different species and groups have a different primer binding site and hence will result in the different amplification site. For example, in this case what you see here is that all these species like S1 to S2 and 10, they are very close because their amplicon size is very small or their amplicon size is similar.

Whereas in this case, when the amplicon are widely distributed, they may not be they may not be closer enough right. And that is why when you do the PCR, you are going to see the amplicons coming from the S11 to S20 different and that can be used very nicely to know which one is the close relative and which one is the distant relative. Then we also can study the PCR for detecting the plant microbe interactions. So, in this particular case, you are going to use the PCR and as well as the RT-PCR techniques and you can be able to detect the different types of pathogens which are associated with the particular crop. For example, the clavibacter which is associated with the potato, then we also have the acidovorous which is actually associated with the watermelon and so on.

So, you can actually be able to take the plant sample and then you actually can do the PCR for these infectious organisms. Then you also can use the PCR in the tissue culture. So, you can actually be able to use the PCR for detecting the different stages of the tissue culture species. So, the use of PCR in tissue culture was already reported in the year of 1992. It was used in the analysis of DNA and a specific gene in the plant cell at the different stages of regeneration during in vitro culture along with the RAPD or the random amplification of the polymeric DNA technology.

The level of polymorphism in the regenerated plant could be revealed by these two technology. PCR could flawlessly amplify the neomesin phosphatase gene and antibiotic which is used for the selective marker into the transgenic animals. And the PCR is also been used in the veterinary parasitology. So, it is been used for detecting the different types of parasites which are been present on to the animals. For example, you can actually be able to use the Ozuski disease virus of the pigs and this can be detected.

Then we can also use the bovine leukemia virus or BLV. This virus causes the zoonotic

bovine, bovine leukosis and PCR assay for detection was developed in the year of 1991. Then we also can use the virus diarrheal virus or the bovine viral diarrhoea virus. This virus is not only fatal to the cattle, but also causes the contamination in the calf serum used in the cell culture work thus leading to the contamination of the vaccine and the pharmaceutical products. Besides the above example, the PCR has been used in the routine diagnosis of the veterinary viruses such as the porcine, parvovirus, bovine, papilloma virus and avian, polymya virus, chicken anemia, duck hepatitis and so on.

So, PCR is having an extreme potential in detecting the different types of diseases. Then PCR can also be used into the forensic science. So, for example, the criminal investigations. So, each individual has a different DNA profile known as the DNA fingerprinting. A DNA fingerprinting uses the variable number of tandem repeat or VNTR loci as these loci is so variable that the unrelated individuals are unlikely to have the same VNTRs which means the VNTRs pattern of a VNTR is a specific to the particular person not to particular family ok.

So, a sample of the blood, hair root or the tissue left in the crime scene can be used to identify a person using the PCR by comparing the DNA of the crime scene with that of the suspect or with the DNA database of the earlier convicts. Evidences from the decade old crime can be tested confirming or defending the people originally convicted. So, you can imagine that you have actually you have collected the crimes in DNA and these are the some of the pattern what you got. So, these are the VNTRs what are present and then you can actually be able to do the similar PCR with the different types of suspects. So, you actually know which are actually going to be the prime suspects of or the suspects which were present at the crime site and then what you can do is you can actually be able to do the analysis.

So, what you see here is that the suspect one is actually having a VNTR pattern which is very different. So, for example, in the crime scene DNA what you got you have the VNTR like 1, 2, 3, 4, 5 and 6. Now 1 is correct. So, if you draw the line you will see that the 1 is present in this one 2 is not present in this one 1 is present in this one and 1 is slightly present in this one also. Then 2 is present in this one, but 2 is absent in this one, but 2 is present in this one. suspect number 3.

new. then the nu3, nu3 is absent in the suspect number 1 nu3 is present in the suspect number number 2 and nu3 is also present in the suspect number 3 and nu3 is also present in the suspect number 4 which is also absent in suspect number 1. So, suspect number 1 will become the innocent because the majority of those bands are not present and suspect number 2 also does not have. So, it is also going to be innocent whereas, the band number 4 is also present in the suspect number 3, band number 5 is also present in the suspect number 5, suspect number 3 and the band number 6 is also present in the suspect number 3. So, this means the this particular person is actually his DNA profile is matching with the crime scene DNA what you have isolated. This means the suspect number 3 is actually a

person who may be involved in this particular type of criminal activity, he may be the convict and he may be the present at the crime site.

So, that actually can be proved and then you can actually be do the further investigations. Then the you can also be able to use the PCR for the parental testing. So, similar kind of approach. So, PCR technology also being in finding the biological parents of the adopted or the kidnapped child where the DNA of a child is matched with the close relative. The actual biological father of a newborn can also be ruled out.

In parental testing the short tandem repeat or STRs are being used as a marker where each person's DNA copy contain the two copy of these markers are often from the father and mother. These markers differ in the length and sometime the sequences. So, DNA marker you can actually be able to have like some of these markers like A, B, C, D, E and the mother and the father will you can test the different types of mothers and you can actually be able to test the different types of probable fathers and mothers and all that. So, in the child the STR what is present in 26 and 30 right and that is the combination of this mother and this father because from the this mother it has taken the 26 and from the father it has taken the 30. So, that is why this is these are the potential father and mother.

Similarly, in the mother we have the 8 and 9 and it has taken the 9 from here right. Similarly, from the here it has taken the 10 actually. So, that is all what you see here is that the child's profile is a combination of mother and father right and that is why this child belongs to this particular mother and father. I am not showing you the more samples or the data from the different samples and so on. So, another sensitive technique that can be used to establish the maternal relationship between the people is called as the mitochondrial DNA analysis which relies on the PCR.

This analysis is better than the fingerprinting for the sample which belongs to too old that are nucleus of the cell get degraded. So, in the some cases you are actually going to analyze the mitochondrial DNA because the mitochondrial DNA remain constant and it runs without getting diluted from the single family. The only difference is that the mitochondrial family mitochondrial DNA analysis will tell you who will be the mother and it is actually going to only follow from the mother side ok. So, it is actually going to tell you who will be the mother, who will be the grandmother, who is going to be the grandmother to mother and so on. So, because of the simple reason that the mitochondrial DNA is always coming from the female side it does not come from the father side.

Then, we also can use the PCR in the research applications. So, PCR can be used for the DNA cloning, PCR can be used for the sequencing, PCR can be used for the sequence tag sites, PCR can be used for the phylogenetic analysis and PCR can be used for the gene expression analysis. So, in the phylogenetic analysis you are going to take the relationship between the different types of organisms. So, this is all about the application of the PCR into the application of the PCR. Now let us move on and we will discuss about the

application of the real time PCR.

So, application of the real time PCR is also very diversified and it is as big as the PCR because it can be used for the gene expression profiling, it can be used in the drug discovery, it can be used for the disease diagnosis and management, the PCR RT-PCR can be used even for the viral quantifications, food testing, GMO food, it can be used for detecting the gene expression as well as the gene copy number and it also can be used for the cancer research. So, let us discuss first about the diagnostic part because that is the major area where the real time PCR is always been used and you might heard about the real time PCR when we were talking about the COVID-19 screening and detect diagnostics. So, let us take an example of the detection of the COVID virus or coronavirus during the COVID-19. So, COVID-19 is an infectious disease caused by the SARS-CoV-2 virus and the common symptom include the high fever, cough and shortness of the birth. So, how to test the SARS-CoV-2? So, what you are going to do is I am sure you might have seen all these steps in television and other places that you are going to collect the sample from the cotton swab or from the otoforangial swab.

So, you are going to collect the sample from the nose or from the mouth. So, that sample will contain the virus and it is going to contain the viral RNA because the COVID is a RNA virus. So, swab is collected in the test tube containing the viral transport media or VTM. Then in the step 2, you are actually going to do you are going to extract the RNA using any of those method what we have discussed you can use actually the RNA extraction method and then you are going to get the purified RNA. Then in the step 3, you are going to connect you are going to you know convert that RNA into the complementary DNA.

So, you are going to make the complementary DNA strands. So, then the CADNA tag with the fluorescence for the performing the real time PCR and then you are going to do a real time PCR with the help of the primers targeting to the COVID virus and then you are going to do the analysis. And what you are going to analyze? You are going to analyze the CT values and it a positive value is actually a CT value beyond a certain copy number is going to tell you that the sample is positive whereas, if the sample is below the baseline then it is going to say the negative. So, this is the step sample is collected from the nasopharyngeal swab and collected in a tube containing the viral transport media. The collected sample are processed to extract the RNA it is the most crucial step in the whole process.

Then obtain the real time PCR kit specifically dedicated for detecting the SARS COVID 2 that include the primer and a probe designed to target the specific region of the viral genome. Then prepare the real time extraction media by adding the extracted RNA primer probe and other necessary component as guided by the manual. Then run the real time PCR this cycle will go to the series of changes in the temperature which lead to the amplification of the viral RNA which it is present in the collection samples. And then the probe which are mixed with the reaction mixture will emit the fluorescence then it binds to the viral RNA and the PCR machine will monitor the fluorescence signal. The real time PCR

machine continuously measure the fluorescence and record the cycle at which the cycle signal crosses the preset threshold and this is called as the threshold cycle or the CT value.

The CT value provide the information about the concentration of the viral RNA present in the collected sample. Lower the CT value high will be the viral concentration higher will be the CT value the lower will be the viral concentrations. A positive result is typically been determined by comparing the CT value with the predetermined the threshold value. If the CT value falls below the threshold value of the sample then it is considered as the SARS COVID 2 or it is going to be called as the positive for the COVID 19. Quality control include the positive and negative control to check for the contamination and to validate the results.

Apart from that you can also do the gene expression analysis. So, I am sure when we were discussing about the real time PCR we have shown you a complete demo how you can be able to do the gene expression profiling. So, this we are not going to discuss in detail like where in this one the real time PCR is a powerful technique to analyze the gene expression and in a biological sample. So, what you are going to do is you are going to do a cDNA synthesis, you are going to do the PCR amplification, you are going to analyze the sample to know what is the threshold value then you are going to do a quantification and ultimately in the step 5 you are going to do the data analysis. And this all we have discussed in detail when we were discussing about the real time PCR both theoretically and as well as the experimentally where the students have shown you how to perform the analysis part, how to perform the cDNA synthesis and so on.

Then you can also be able to do the gene copy number. So, utilizing the similar kind of approach you can also be able to do the gene copy number and you can be able to detect the gene copy number utilizing the real time PCR. Then we can also be able to use the real time PCR for detecting the GMOs what is present in the particular or GMOs present in the particular sample and there are multiple ways and in which you can be able to do that. So, in this one what you are going to do is you are going to collect the sample right which is including any GMO sequences and then you design the gene primer which is a specific to the selected gene sequence that are unique to the GMO of the interest. Then you are going to perform the real time PCR reactions and then you are going to do the amplification and detection. Ultimately you are going to know the CT values and if there will be a CT value beyond the threshold then it is actually going to say that the particular genetically modified organism sequences are present in your DNA.

Then you can also use the real time PCR for the cancer cell detections. So, quantifying the cancer cell using the real time PCR involves detection and measurement of a specific gene marker associated with the cancer and this is been published in one of the such article which is called as BMC cancer where they have used the gene expression profiling of circulating tumor cell in the breast cancer by the RT-PCR. So, here also you are going to take the sample you are going to collect the samples then you are going to do the gene marker

selection. So, you are going to select the gene which is specific only for the cancer cells then you are going to design the primers and probes and all that and then you are going to perform the real time reactions then you are going to analyze the data and you are going to say whether the CT value is above the threshold or not and then based on this data you can be able to say whether the particular sample has a cancer cell or not. So and then real time PCR is also been used for the drug discovery.

So, it can actually be able to use for the target identification, pharmacogenomics, biomarker analysis, gene expression profiling, toxicity assessment, pharmacokinetics and pharmacodynamics and assay development and as well as the high-to-fruit screening. So, this is very extensively been you know exploited the real time PCR and as well as the PCR for detect for doing the many types of you know applications. And what we have discussed we have discussed only the application which are actually very very popular right for example, the detection of the COVID and all those kind of things, but that is not the limit of these techniques. These techniques are having the enormous potentials to be used and they can be exploited in a multiple ways in which to know the particular type of. So, apart from the their contribution in the clinical practices or diagnostics purposes they are also having the enormous applications in the basic sciences and so on.

So, all those we have not discussed in detail because then it becomes a very very extensive lecture and it may not be possible to discuss all those in a single lecture or in this particular course. So, with this I would like to conclude my lecture here in our subsequent lecture we are going to discuss about the application of the genome editing and how the genome editing tools what we have discussed in the previous module and how they can be used and exploited for the human welfare. So, with this I would like to conclude my lecture here in a subsequent lecture we are going to discuss some more applications of the molecular biology. Thank you. .

Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 12

Applications of Molecular Biology

Lecture-52 Applications of Molecular Biology (Part 4)

Hello everyone, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati. Today we are going to discuss about molecular biology. We are going to discuss about the different applications of the techniques that have been involved during the process of the molecular biology. So, there are molecular biology principles which are being utilized to develop the different types of diseases or different types of applications.

These applications are enormous. There are multiple ways in which the molecular biology can be exploited. What we were discussing or what we have focused in this particular course, we have focused only on to the four aspects. What we have discussed? We have discussed about the application of the PCR in genetic engineering because the PCR is the base in which the which is being utilized to develop the genetically modified organisms or which are being basic of the developing the genetic engineering or performing the genetic engineering task.

Then we also discuss about the how the transgenic animals which are being produced after the molecular biology is being utilized and then we also discuss about the PCR based applications. In the previous lecture, we have discussed about the PCR based applications where we have discussed about the real time PCR, we have discussed about the PCR and so on. In today's lecture, we are going to discuss about the genome editing tools and their applications. So, what we have discussed? We have discussed about the genetic engineering applications or applications downstream to the genetic engineering where we have discussed about how the genetic engineering is being led to the production of the different types of proteins, enzymes and all other kinds of things and how they are being utilized for the human welfare. And they and genetic engineering is also being used for producing the different types of transgenic animals.

And these transgenic animals are actually serving the multiple purposes. They are being served as a disease model, they are being utilized for producing a particular type of protein which is actually causing the less allergic reactions and then so on. And in the previous lecture, if you recall we have discussed about the PCR based applications both in terms of the diagnostics and as well as in terms of the their contribution into the basic sciences. So,

the PCR and as well as the RT-PCR or real time PCR is extensively being used in answering the different types of questions. Now, in today's lecture, we will discuss about the genome editing and how the genome editing tools, what we have discussed in the previous module can be exploited for the deployment of or for making a genetically modified or making or how the genetically genome editing tools can be used for editing the genome and that in turns will actually be good for the human welfare.

So, let us first start with the so, we if you recall in the previous module, we have discussed about the three genome editing tools. We have discussed about the zinc finger nucleus, we have discussed about the talon and the we also discuss about the CRISPR cache. And all of these are actually having their own positive and negative. And we are not discussing about the homologous recombinations and as well as the non homologous recombination as a tool. What we are discussing? We are discussing about the tools what are been currently been utilized or what are the tools which are you know modern tools which are being used for the genome editing.

So, we are you know first start with the zinc finger nucleus, then talon and the CRISPR cache. So, talon the major application of the talon is that it is been used for the gene knockout, gene knock ins, targeted chromosomal deletions, modification of the genome, crop improvement and the application in the HIV treatment. So, let us first discuss about the gene knockout and the knock ins. So, the knockout involves the inactivation of the gene or the gene segment resulting in the loss of gene function. This is typically achieved by introducing the random insertion or the deletion into the genome to the imprecise DNA repair process of non homologous end joining or NHEG targeted by the double stranded breaks.

Before the advent of CRISPR class IX, researchers commonly rely on the random mutagenesis methods such as ENU mutagenesis induced by the radiation or the chemical for generating the DNA changes. This approach often employed in a zebra fish towards skinning was less precise. In contrast, the knock ins operate differently by introducing the specific base pair changes with the precision into the genome. This involves the targeted introduction of a template such as the synthetic DNA fragment designed to replace an existing genome sequence. The highly accurate repair process known as the homologous directed repair HDR guides these precise changes making a departure from the random distribution characteristic of the knockout models.

The process of genetic modification often initiated with the introduction of the double stranded breaks in the DNA and efficiency of precise DNA modification can be significantly improved by generating a site specific DNA. Zinc finger nucleus is served as a powerful tool for inducing the specific DNA cleavage since the variable DNA binding domain of zinc finger nucleus can be customized to recognize a specific DNA sequence chosen by the investigator. By manipulating the outcome of the DNA repair through the different selection method, it becomes positive possible to achieve either the gene knockout or the precise insertion of

the targeted transgene or the knock ins. The mechanism of the gene knockouts and the gene knock ins or the knockout by the zinc finger nucleus involves their ability to introduce induce that the targeted double strand breaks in the DNA and exploit the cellular repair machinery. Here is a breakdown of the process.

In a gene knockout, you are going to disrupt a particular gene. In a gene knockout, the objective is to disrupt or eliminate the expression of a particular gene and zinc finger nucleus are designed to target and bind to the specific DNA sequence of the gene to be knocked out. So, you are going to do the double stranded breaks and the zinc finger nucleus induce TFN at the target gene locus activating the cells repair mechanisms. Non-homologous end joining or NHEG in the absence of template for the homologous directed repair, the cell often employed the error prone non-homologous end joining pathway to repair the double standard break. This frequently result in the small insertion or deletion at the side of the break.

So, what will happen is that you are actually going to use a zinc finger nucleus. It is actually going to cut a specific DNA sequences and then it is actually going to generate the double standard breaks and those double standard breaks are actually going to be filled with the help of the exploiting the non-homologous end joining pathway. And then ultimately there will be a functional disruption. The Intel introduced during NHEA can lead to the frame sheet mutations, premature stop codon or other alteration that disrupt the normal functioning of the target gene resulting in a knockout phenotype. So, this is exactly what is going to happen that zinc finger nucleus what you have introduced is actually going to recognize a specific DNA sequence and that is how it is actually going to introduce our double standard DNA break.

And then these double standard DNA breaks are going to be filled with a random sequence. And as a result of these random sequences, it is actually going to cause the frame sheet mutations, it is going to cause the deletions and other kinds of things and as a result it is actually going to make this particular gene non-functional and that is how it is actually going to have the gene knockouts. This means that particular gene is no longer be functional in this particular in this organism containing this particular genome. And similar to that we also can do the gene knock ins. So, introduction of a desired sequence to achieve the gene knock in a specific sequence is desired to be inserted into them. This can be done by a new gene or modified version of an existing gene.

Then you are going to design the zinc finger nucleus. So, zinc finger nucleus are engineered with a zinc finger DNA binding domain designed to recognize and bind to the target side of the genome where the insertion is intended. Then you are going to generate the double standard breaks. So, zinc finger induce with double standard breaks precisely at the target side and then you are going to do the homologous directed repair mechanisms and as a result the homologous sequence are actually going to replaced the this double standard break DNA and that is how you are actually going to introduce the new gene into

the organism or by the genome. So, what you are going to do is first step is same that you are actually going to use the zinc finger nucleus it is going to break cause the double standard break and after that you are actually going to put a donor DNA right.

This donor DNA is going to have the homologous arm right and this homologous arm is actually going to go with the homologous recombination and as a result this particular DNA is actually going to be introduced or going to be inserted into the genome and that is how it is actually going to be resulted into the gene knock ins which means you have actually introduced a new gene into the organisms or the genome. Now you if since these two processes can be possible that you can actually be able to remove a particular gene you can be able to introduce a particular gene you can also be able to do the mutagenesis right. So, then you can actually be able to do the insertion of a mutated gene you can actually be able to do other kinds of things. So, that also can be a application that you actually can use the zinc finger nucleases for the targeted mutagenesis. So, in that case what you are going to do is you are going to just remove a particular nucleotide or a codon and then you can actually be able to put your specific codon what you want right.

Then you can also use the modification of the genome with the help of the zinc finger nucleus. So, just what we have discussed so far right you can actually be able to do the knock in or knock out and all those kind of thing and as a result you can be able to even modify a genome using the zinc finger nucleus. This is one of the example what is showing right there is a pig right and it actually has the normal cell and what you can do is these you can just isolate these normal cells which actually contains you know the wild type genome. So, I will say the wild type genome right and all the cells are actually containing the wild type genome right what you can do is you can utilizing the zinc finger nucleus you can actually be able to introduce or target some of the cells and you can actually be able to generate the for example, the mutated genome right. So, you can actually be able to make the further genome and then you can actually just isolate these cells and then you can just introduce these cell back into the pig and what you can see here is this pig does not have a skin pattern, but this pig is actually showing a skin pattern because you have introduced a gene which will actually cause the patterning onto the skin and that is how it is actually been cause the production of the transgenic pig actually right.

You might have seen that I have I have discussed different types of transgenic animals and this is the more or less the smarter way of generating the transgenic animal. Similarly you can actually be able to do many things like you can actually be able to do the multiple ways in which you can be able to generate the transgenic animals. You can also be able to use the deletion of a chromosomal sequences in the human cell using the zinc finger nucleus and this is what it is actually going to say that you are actually going to target a particular DNA sequence and that is how you can be able to remove the particular sequence and you can actually be able to put the new sequence and that is how you can be able to generate deletion and as well as the addition into the human cell using the zinc finger nucleus. Since you can be able to do that you can actually be able to use that for crop improvement. So

advancement of the horticulture crop with desired trait in the contemporary and future agitators rely significantly on the genetic engineering.

Reverse genetics has played a pivotal role in cropping improving the crop genetic quality by elucidating the gene functions and that is how you are actually going to use the multiple approaches. You can actually be able to use the zinc finger nucleases you can actually be able to use the targeted delivery of that particular sequence and that is how you can actually be able to use the agrobacterium mediated DNA transfer and so on and that is how you can be able to make the genetically modified crops and that will actually going to improve the crop yield and other kinds of things. Then another crucial application of the zinc finger nucleus is the in the plant biotechnology is the development of the stacked transgenic traits. Currently the gene stacking process involves the breeding stacks of the randomly inserted event requiring the independent sorting characterization and intercession into the elite germplasm. This method is both time consuming and expensive.

Zinc finger nucleus offer a solution by precisely integrating a new transgene into a previously integrated trade land pad facilitating the creation of the transgenic trade stacks. In this way the herbicide resistance traits were stacked using the precise precision gene targeting through the gene finger nucleus. So, this is what it is actually going to say that you are actually going to target a particular sequence into the gene and that is how you are actually going to generate or put the particular gene and that is how you can actually be able to make the herbicide resistance crops and so on. Then we can also be able to use the chromosome with the NMC targets and so on and this is more or less they all are actually going to have the similar kind of application where you are going to use the cut of the particular genome then you are going to introduce the new DNA and so on and you can actually be able to use the zinc finger nucleases for that particular applications. Now, let us move on to the next tool and the next tool is the TALENs.

So, application of the TALEN. So, customized zinc finger nucleus and TALENs have revolutionized the genetic research by efficiently introducing the target state alteration in the various organisms and cell type. And the application wise the TALEN and the zinc finger nucleus are doing exactly the same except that the tools are different. Zinc finger nucleases are little difficult to perform compared to the TALEN, but the overall mechanisms remains the same. You are actually going to have you can be able to do all the things what you are doing with the zinc finger nucleus with the help of the TALEN. So, in the help of the TALEN also you are actually going to produce the like for example, you can actually be able to produce the disease resistance rice.

So, this is an example of how you can be able to use the reduction of the disease resistance rice right. And here what we have they have done is they have just used the TALEN to introduce a double standard break and then they have introduced a new gene which is called as OS1193 gene and that is how they are actually been able to produce the disease resistance rice. Similarly, you can be able to use the mouse editing as well as with the help

of the TALEN based gene editing. So, the Y chromosome which is a unique structure and specific gene contain crucial for the male sex determination and fertility presents challenge for the conventional gene targeting in the mouse embryonic stem cells. So, previous attempt using the traditional studies have been largely unsuccessful in generating mutation in the Y link gene, but with the help of the TALEN you can be able to even generate the mutations into the Y chromosome.

And then you can actually be able to use the TALEN based knockout as well as the knock-ins right. So, you can actually be able to use the TALENs offer a versatile and accessible platform for modifying the livestock genomes. And in contrast with zinc finger nuclease the TALEN overcome the technical challenges and the budget constraint. So, that is only the differences between the TALEN and the zinc finger because zinc finger is slightly costly and it is more challenging to perform compared to that the TALENs are easy to perform and they are actually more cost effective. So, then TALEN activity in the bovine embryo.

So, these are also going to be one of the application area where the TALENs are being used for you know developing the different types of embryos and all the kinds of things. Then TALENs are also being used for genome editing in protecting the you know the cells or protecting the organism from the HIV infections. So, HIV primarily utilizes the coreceptor like CCR5 and CXCR4 along with the CD4 for cell entry. So, individual with a homozygous CCR5 delta 32 deletion exhibits resistance to the HIV. So, if you have a CCR5 which does not contain the 32 right which does not contain the 32 region then it is actually going to resistance for HIV.

So, it is like CCR4 silencing or anti CCR antibodies and small molecule inhibitors have shown efficacy in inhibiting the HIV infection. So, genome editing approach initially focused on the CCR coreceptor using the zinc finger nuclease progress to the human trial with the phase 1 clinical trial showing promises. So, the CCR4 concern about off target effect and cytotoxicity remain with the zinc finger. TALEN offer advantage in terms of minimal cytotoxicity and off target compared to the other genome reading technology. So, with the help of that you can be able to target the CCR5, you can be able to target the CXCR4 and so on and that is how you can be able to develop an organism which is HIV resistance.

TALEN have emerged as a powerful genome editing tool with a wide range of lecture. One notable application is the modification of a model organism where TALEN enable the efficient introduction of the targeted alterations. So, there are enormous applications of these genome editing tools and we cannot you know and there are so many different avenues which are still been exploring by the scientist. So, since the applications are endless and the potential of these tools are enormous it is difficult to encompass those into the single lecture ok. So, what we have discussed? We have discussed about the application of the zinc finger nuclease, the mechanism of how the zinc finger nuclease are causing the knockout or the knock ins and then we also discuss about the application of the TALENs.

So, with this I would like to conclude my lecture here and in subsequent lecture we may discuss some more aspects related to application of the molecular biology. So, with this I would like to conclude my lecture here. Thank you. Thank you.

**THIS BOOK
IS NOT FOR
SALE
NOR COMMERCIAL USE**



(044) 2257 5905/08



nptel.ac.in



swayam.gov.in