Biochemical Engineering

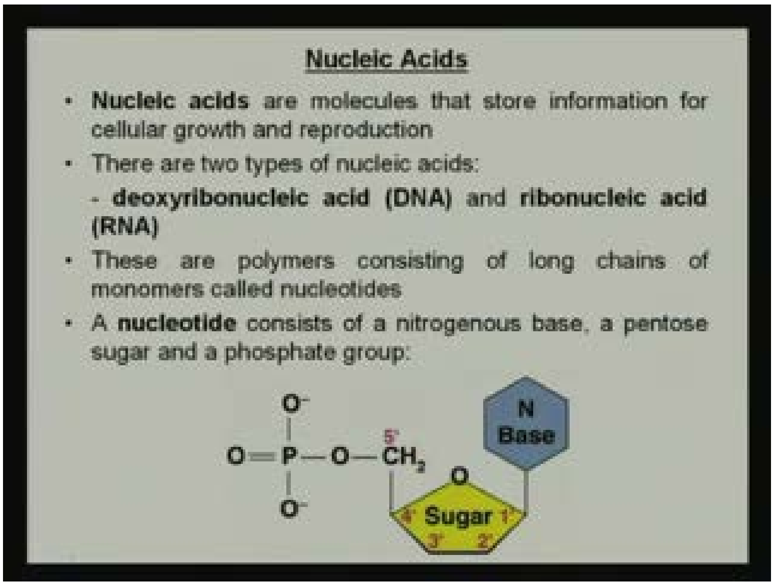
Prof. Dr. Rintu Banerjee
Department of Agricultural and Food Engineering
Prof. Dr. Saikat Chakraborty

Department of Chemical Engineering
Indian Institute of Technology, Kharagpur
Module No. # 01
Lecture No. # 05

Nucleic Acid

Good morning students. In my earlier classes, I have already discussed about some of the biological macromolecules which are present in the cell. Those molecules are either protein, carbohydrates and so on. Today I will discuss on another biological macromolecules which are present in the cell called nucleic acids.

(Refer Slide Time: 00:55)



Nucleic Acids

- **Nucleic acids** are molecules that store information for cellular growth and reproduction
- There are two types of nucleic acids:
  - **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**
- These are polymers consisting of long chains of monomers called nucleotides
- A **nucleotide** consists of a nitrogenous base, a pentose sugar and a phosphate group:

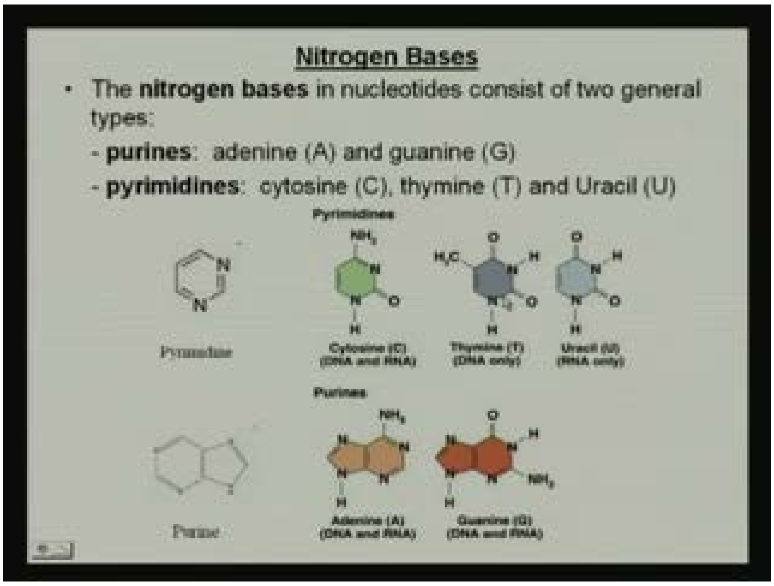
Nucleic acids are the molecules that store information for cellular growth and reproduction. If we see broadly classify the entire nucleic acids, it can be divided distinctively into two major groups. One is called deoxyribonucleic acid or D N A, and ribonucleic acid or the R N A. These are the polymers consisting of long chains of monomers called the nucleotide.

So, what is nucleotide? A nucleotide consists of a nitrogenous base, a pentose sugar, and a phosphate group. That means in other word, we can tell that in a particular nucleotide, we have a nitrogenous base, one sugar moieties and one phosphate group is there.

Now, this is the structure of a nucleotide. Say this is a pentose sugar. In carbohydrate classes I have already discussed that what is pentose? What is hexose? So, this is a pentose sugar. If we number this carbon then this is the first number of carbon, second, third, fourth and fifth. So, this is a pentose sugar and the first carbon is being occupied by a nitrogenous base.

So, here one nitrogen is attached to the sugar moieties, and to this fifth carbon of this particular sugar one phosphate group is attached. And when this sugar nitrogen and phosphate groups are there. Then we can call it as nucleotide.

(Refer Slide Time: 03:02)



Nitrogen Bases

- The **nitrogen bases** in nucleotides consist of two general types:
  - **purines**: adenine (A) and guanine (G)
  - **pyrimidines**: cytosine (C), thymine (T) and Uracil (U)

Now, what is this nitrogenous base? Basis which are there along with this particular nucleotide molecules . So if we see this nitrogenous basis in nucleotide, it consists of two general type .One is the purine and other is the pyrimidine. Now if we see the structure of purine and pyrimidine .Then we can find that this pyrimidine is a nitro containing heterocyclic molecule.

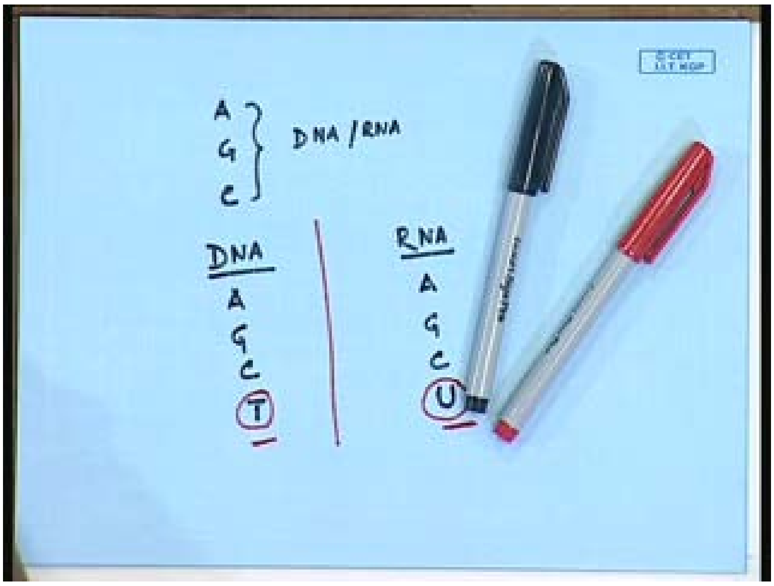
Now,here if we see further, this pyrimidine. It has got three main derivatives, one is called cytosine, in it is symbolized as C, Thymine t and uracil U. When to this heterocyclic nitro containing ring, one imidazole group is being attached, it form the purine. And, if we see that derivatives of this purine, we can find that adenine and guanine. That adenine is symbolized as A and guanine is symbolized as G. These are the derivatives of purine.

Now, if we further go into this derivatives of purine and pyrimidine. We can find that this adenine and guanine. This A and G they are the purine derivatives are present both are present in R N A as well as D N A molecules. Whereas, if we see the pyrimidines and pyrimidines derivatives, then we can find that, this cytosine is present both in D N A and R N A .Whereas, thymine is present all the in D N A and uracil is present only in R N A.

That means structurally, if we see the structure of this pyrimidine derivatives. That is thymine and uracil, we can find that see rest of their structures are same. Both the molecules except this methyl group which is attached to this thymine molecule, and these thymine molecules. When this methyl group is removed or this methyl group is not attached, we can call it as uracil .That means in other word, we can tell that this uracil methyl containing uracil is the thymine or demethylated thymine is the uracil.

So, these are the structural differences and here this thymine is present in D N A. And in R N A thymine is substituted with uracil, that is U. That means in short, what we can conclude from this nitrogenous base?.

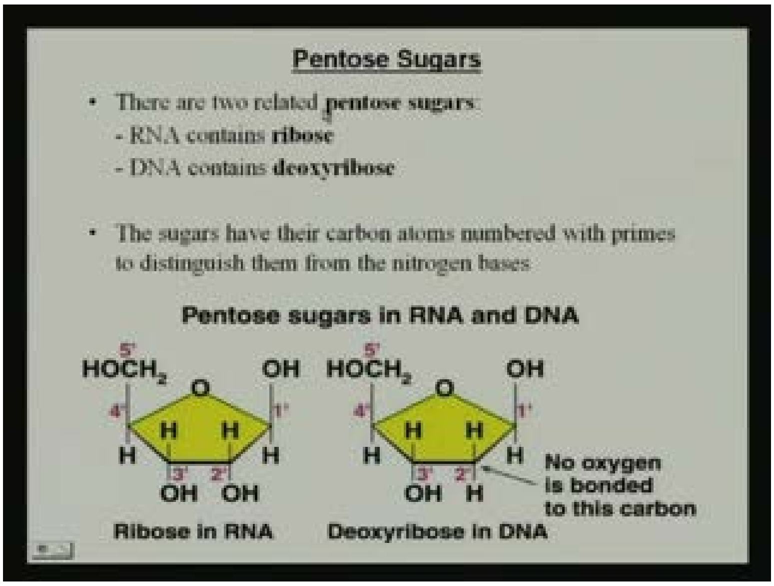
(Refer Slide Time: 06:42)



We can tell that A G C is common, that means adenine, guanine, cytosine is common in both D N A as well as R N A molecules. Whereas, in case of D N A, when we are considering D N A ,then we can tell A G C and T .Where as in case of R N A ,when we are talking about R N A, we can write A G C and U. This U, this thymine is changed with uracil, this is the difference between one D N A molecule and one R N A molecule.

So, these are all about the nitrogenous bases. And these nitrogenous bases are attached to the first carbon of the pentose sugar. So, then we are coming to this sugar moieties.

(Refer Slide Time: 07:50)



Pentose Sugars

- There are two related **pentose sugars**:
  - RNA contains **ribose**
  - DNA contains **deoxyribose**
- The sugars have their carbon atoms numbered with primes to distinguish them from the nitrogen bases

Pentose sugars in RNA and DNA

No oxygen is bonded to this carbon

Ribose in RNA

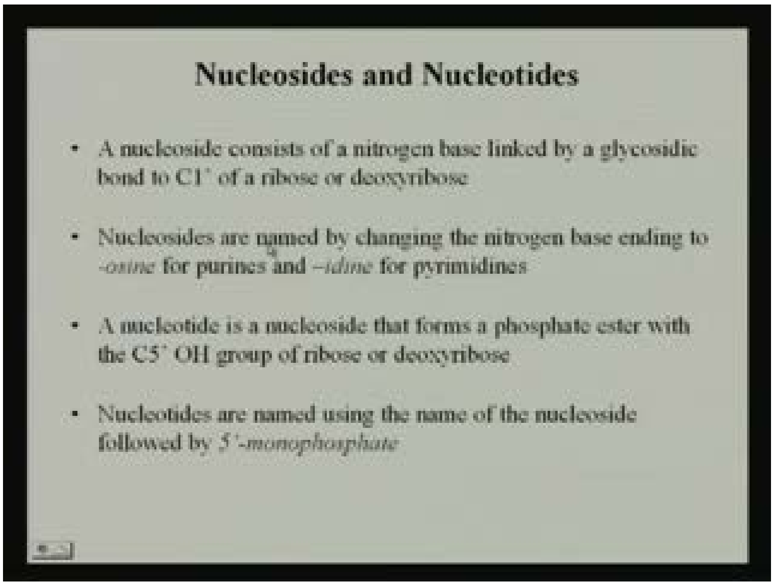
Deoxyribose in DNA

.

Now, as I had told you that five carbon sugar moieties, and that is why? We are talking, calling it as pentose sugar. Now this sugar, these pentose sugars are also of two types. In case of D N A as well as R N A .In case of R N A, this is the simple ribose sugar, this is the ribose sugar ,see this is 1 2 3 4 and 5 carbon sugar .And this is the ribose sugar in case of R N A molecule. But, here in case of D N A it is called deoxyribose. That means one oxygen is missing in the sugar moieties, five carbon sugars.

So, what is that see in the second carbon? Here you see one H and O H is there. But, here in this deoxyribose, you see oxygen is missing. No oxygen is bonded to this carbon atom and that is the reason, why it is called deoxyribose? And which is present in the D N A molecule. The sugars had their carbon atoms numbered with the prime to distinguish them from the nitrogenous bases. And to these particular carbon atom number, one here the nitrogenous bases are being attached .And this is the major difference with the ribose sugar five carbon sugar with D N A and R N A.

(Refer Slide Time: 09:46)



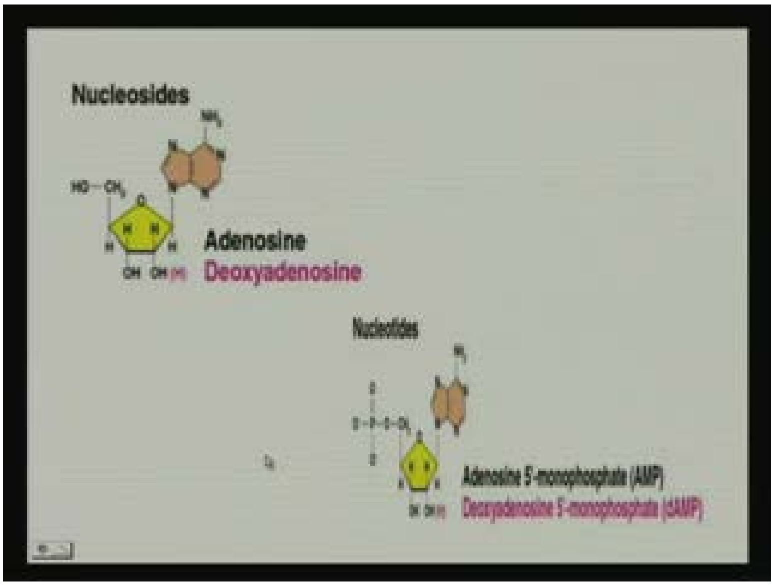
Nucleosides and Nucleotides

- A nucleoside consists of a nitrogen base linked by a glycosidic bond to C1' of a ribose or deoxyribose
- Nucleosides are named by changing the nitrogen base ending to *-osine* for purines and *–idine* for pyrimidines
- A nucleotide is a nucleoside that forms a phosphate ester with the C5' OH group of ribose or deoxyribose
- Nucleotides are named using the name of the nucleoside followed by *5'-monophosphate*

Now,coming to this nucleoside and nucleotide. A nucleoside consists of a nitrogenous base linked by a glycosidic bond to carbon atom number one of the ribose or deoxyribose.

So irrespective of the ribose sugar to the first carbon, these nitrogenous bases are attached to both in D N A as well as R N A molecules. Nucleosides are named by changing the nitrogen base ending to osine for purine and idine for pyrimidines. That nucleotide is a nucleoside that forms a phosphate ester with the carbon atom number 5 hydroxyl group of ribose or deoxyribose. Nucleotides are named using the name of the nucleoside, followed by 5 prime mono phosphates.

(Refer Slide Time: 10:59)



Nucleosides

Adenosine

Deoxyadenosine

Nucleotides

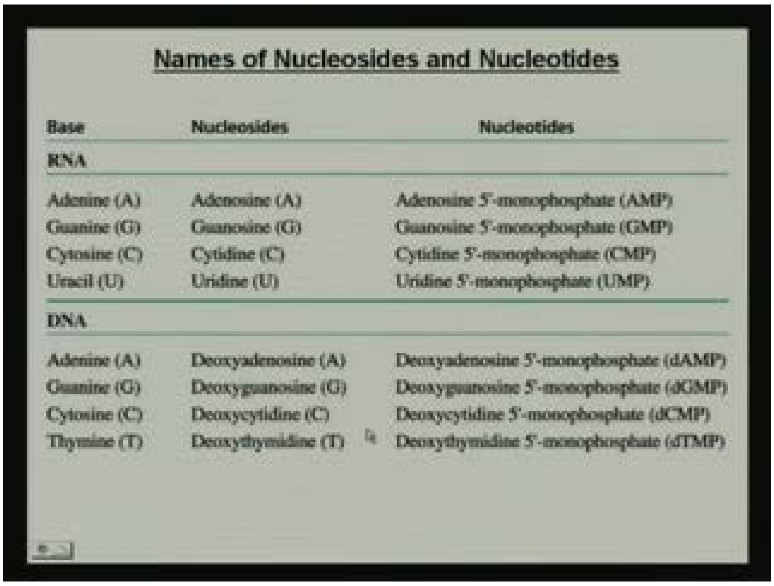
Adenosine 5'-monophosphate (AMP)

Deoxyadenosine 5'-monophosphate (dAMP)

So, if we see the actual structure, then we can find that when we talk about this nucleoside. Nucleoside means ribose sugar, and to this first carbon ,this nitrogen base is attached .So, here you see this is the adenine .Adenine group is attached to this and here if it is the deoxyribose .Then this O will be missing H H then it is the adenosine group is there. In case of nucleotide, see nitrogen and sugar this is nucleoside. When we are talking about a nucleotide then I have already shown you that nitrogen, sugar and to that one phosphate is attached .See this structure, this is nitrogen, sugar through this phosphate moieties are attached to this.

So, it is called adenosine 5 prime mono phosphate. That means it is the adenosine 5 prime mono phosphate, that means to the 5 carbon atom, this phosphate group is attached .So mono phosphate 5 only one phosphate group is attached.

(Refer Slide Time: 12:30)



Names of Nucleosides and Nucleotides

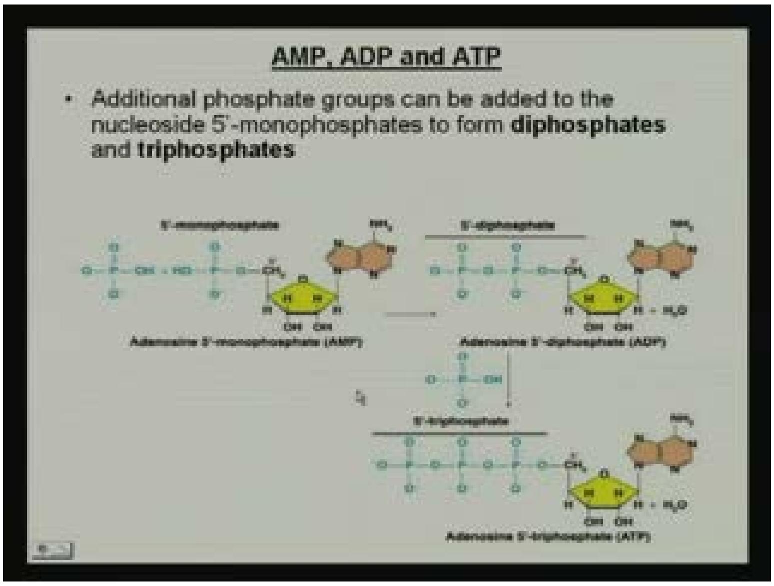
Base	Nucleosides	Nucleotides
RNA		
Adenine (A)	Adenosine (A)	Adenosine 5′-monophosphate (AMP)
Guanine (G)	Guanosine (G)	Guanosine 5′-monophosphate (GMP)
Cytosine (C)	Cytidine (C)	Cytidine 5′-monophosphate (CMP)
Uracil (U)	Uridine (U)	Uridine 5′-monophosphate (UMP)
DNA		
Adenine (A)	Deoxyadenosine (A)	Deoxyadenosine 5′-monophosphate (dAMP)
Guanine (G)	Deoxyguanosine (G)	Deoxyguanosine 5′-monophosphate (dGMP)
Cytosine (C)	Deoxycytidine (C)	Deoxycytidine 5′-monophosphate (dCMP)
Thymine (T)	Deoxythymidine (T)	Deoxythymidine 5′-monophosphate (dTMP)

So, when we are naming this nucleoside and nucleotide .Then how we are naming? In case of R N A ,I have already told you that here the basis are this nitrogenous basis are A G C and U so see A G C and U .This nitrogenous, 4 nitrogenous basis are there in case of nucleoside. We are calling this adenine, guanine, cytosine and uracil as adenosine guanonsine, cytidine and uridine. So, these are the nucleoside naming.

When, we are talking about the nucleotide, then adenosine 5 prime mono phosphate. As I have already mentioned that here nucleotide are named using the name of the nucleoside followed by 5 prime monophosphate. So, similarly, adenosine 5 prime mono phosphate, guanosine 5 prime mono phosphate, cytidine 5 prime mono phosphate, uridine 5 prime mono phosphate and so on.

In case of this D N A, here it is deoxyribose, so obviously deoxyadenosine, deoxyguanosine, deoxycytidine and so on. When we are naming the nucleotide then deoxyadenosine 5 prime mono phosphate or d A M P, deoxyguanosine 5 prime mono phosphate deoxycytidine 5 prime mono phosphate ,deoxytymidine 5 prime monophosphate .So these are the names based on the type of molecules we are discussed..

(Refer Slide Time: 14:370



AMP, ADP and ATP

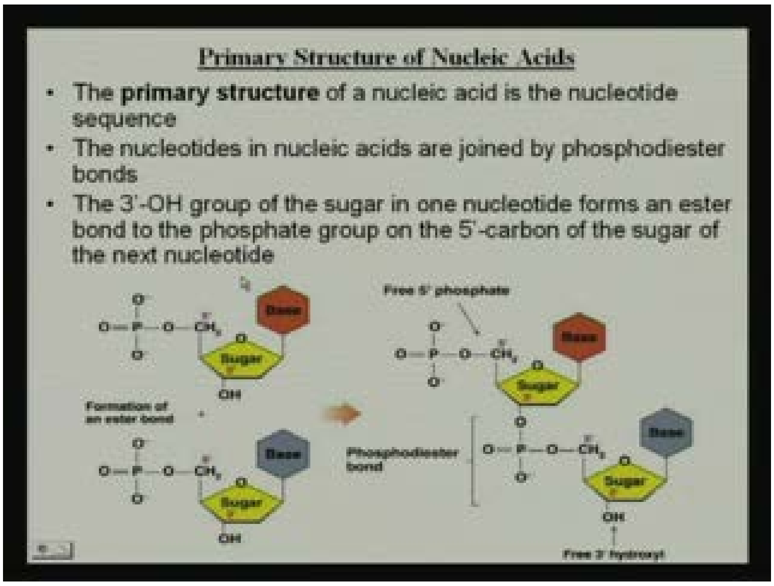
- Additional phosphate groups can be added to the nucleoside 5'-monophosphates to form **diphosphates** and **triphosphates**

So, these are the nucleotide and nucleoside and differences between this nucleotide and nucleoside.

Now, when we are talking about this monophosphate, diphosphate, triphosphate and so on. Then this additional phosphate group can be added to the nucleoside 5 prime monophosphate to form diphosphate ,triphosphates and so on. What does it mean? See this is a nucleotide, nitrogenous base, sugar, and one phosphate is there. If we want to attach another phosphate group to this, that means here, if we want to add upon another phosphate group, then it can be attached.

And here, it forms the adenosine diphosphate. To this diphosphate when we are introduced it to incorporate another phosphate group to this then ,it forms the adenosine triphosphate or A T P molecule .And as, we all know we have already learned that A T P is the major energy source for all cellular activities. So these A T P molecules are playing a significant role in different metabolic activities, which are performed by the cell.

(Refer Slide Time: 16:16)



Primary Structure of Nucleic Acids

- The **primary structure** of a nucleic acid is the nucleotide sequence
- The nucleotides in nucleic acids are joined by phosphodiester bonds
- The 3'-OH group of the sugar in one nucleotide forms an ester bond to the phosphate group on the 5'-carbon of the sugar of the next nucleotide

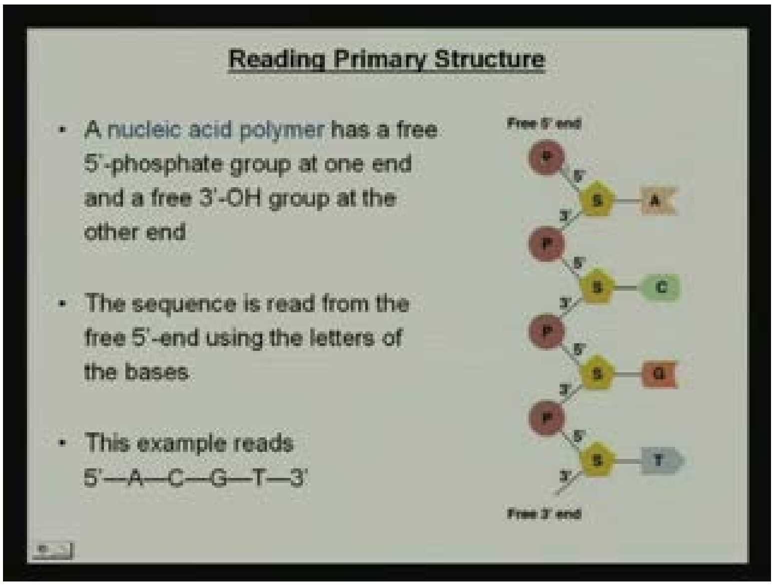
Now, coming to this primary structure of the nucleic acid .The primary structure of the nucleic acid is the nucleotide sequence, the nucleotides in nucleic acid are joined by phosphodiester bonds. The 3 prime hydroxyl groups of the sugar in one nucleotide form an ester bond to the phosphate group on the 5 prime carbon of the sugar of the next nucleotide. What does it mean? It is see this see this is one nucleotide this is another nucleotide.

So, just till now, what we were discussing? That what is nucleoside? And what is nucleotide? We have learned nucleoside is sugar plus nitrogen base. And when one phosphate is being attached to this, it forms the nucleotide. That means in case of nucleoside, no phosphate is attached. In case of nucleotide nitrogenous base sugar moieties plus phosphate. That means sugar plus nitrogenous nitrogen base is the nucleoside. Sugar, nitrogenous base and phosphate is the nucleotide.

Now, when one such nucleotide is being attached to another nucleotide at then this ladder of this linkage or this chain like formation is taking place. Now how this chain is being form, now here you see this is the carbon one two three four five first carbon is already occupied by this nitrogenous base. And third carbon, in the third carbon this phosphate group of another nucleotide is going and it is getting bonded with the phosphodiester bond.

Now, when this type of linkages are taking place, then one nucleotide is getting attached to the another nucleotide. And this phosphate, when it is coming to that next nucleotide that is, third carbon sugar moieties of the another nucleotide. Then it is getting linked and in this way the chain is formed, and this chain is called the primary structure of the nucleic acid.

(Refer Slide Time: 19:07)



Reading Primary Structure

- A nucleic acid polymer has a free 5'-phosphate group at one end and a free 3'-OH group at the other end
- The sequence is read from the free 5'-end using the letters of the bases
- This example reads 5'—A—C—G—T—3'

I think,if I show you in little elaborate way, see this is the sugar moieties this is the third carbon, this is the fifth carbon of the sugar and the first carbon it is the nitrogenous base. This is the fifth carbon, to this fifth carbon phosphate group is attached. This phosphate is once again getting joined to the third carbon of the second nucleotide .Here nitrogenous base is there, fifth carbon here phosphate is there, and this phosphate is being attached to the third nucleotide and sugar moieties.

And third carbon of the third nucleotide, and here from this phosphate is being attached to the fourth nucleotide, molecule the sugar molecule the third carbon of this forth number of this nucleotide .And in this way this 1 2 3 4 and that way nucleotides are being linked up and they formed a chain like structure this nitrogenous bases may be A T

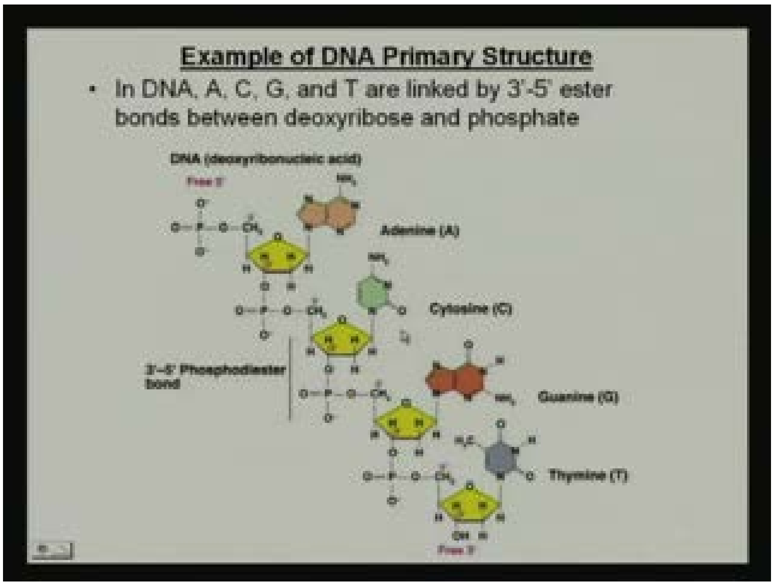
G C, A U, G C so see here A C, G T. These are the nitrogenous bases, so a nucleic acid polymer has a free 5 prime phosphate group at one end. And a free 3 prime, H group at the other end the sequence is read from the free 5 prime end using the later of these bases.

See for example, 5 prime a C G T 3 prime, that means, if we want to write the structure that how this nitrogenous basis are attached? Then we will have to write 5 prime a C G T and then ending with the 3 prime .So this is the example of the primary structure of the nucleic acid. You see this particular linkage has got very close similarity with the primary structure of the protein molecule. But, I have already discussed that one amino acid is being attached to it, another amino acid .The second amino acid is getting attached with third, fourth, fifth and so on.

What is happening in case of these amino acids? This when they form this protein structure, that in one particular amino acid, we have already learned that there is a amino terminal and one carboxyl terminal of that particular two groups are there. So when one amino group is getting attached with the carboxyl group of another amino acid, they form the peptide linkage. And this peptide linkage joins that one amino acid to another and that way the primary structure of protein is formed.

Here, the similar thing is happening with this particular primary structure, where you can see that see that one nucleotide is getting attached with another nucleotide with the third carbon of this pentose sugar. And the phosphate group is being linked up and the chain is getting increased. And that way primary structure of this nucleic acid is form.

(Refer Slide Time: 23:14)

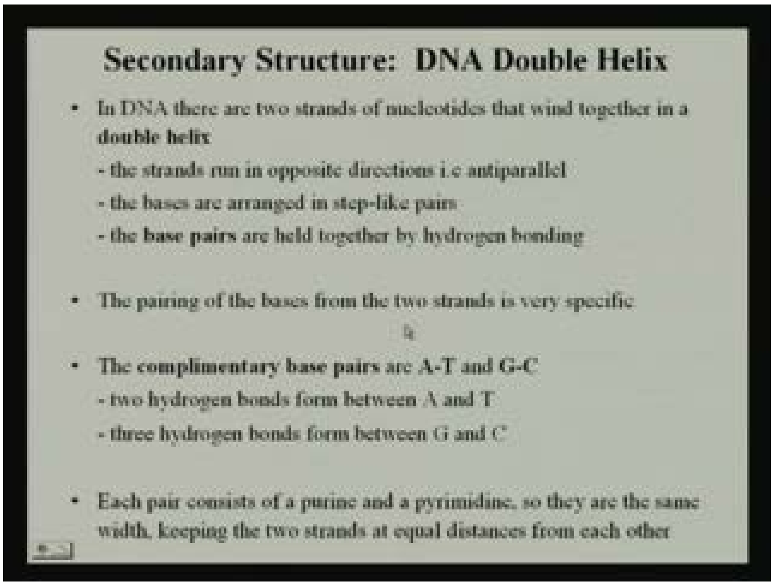


Example of DNA Primary Structure

- In DNA, A, C, G, and T are linked by 3'-5' ester bonds between deoxyribose and phosphate

Now, here this is the example of this R N A primary structure. Here this T is replaced by this U, and here this 3 to 5 prime ester linkages are there. So, here if we want to tell then we can tell this that 5 prime a C G U and 3 prime. Where as in case of D N A you see U is replaced with the thymine. So, if we want to tell the structure of this primary structure of this particular this nucleic acid. Then we can file that 5 prime a C G T and 3 prime so these are the primary structures of this D N A and the R N A molecules.

(Refer Slide Time: 24:14)



Secondary Structure: DNA Double Helix

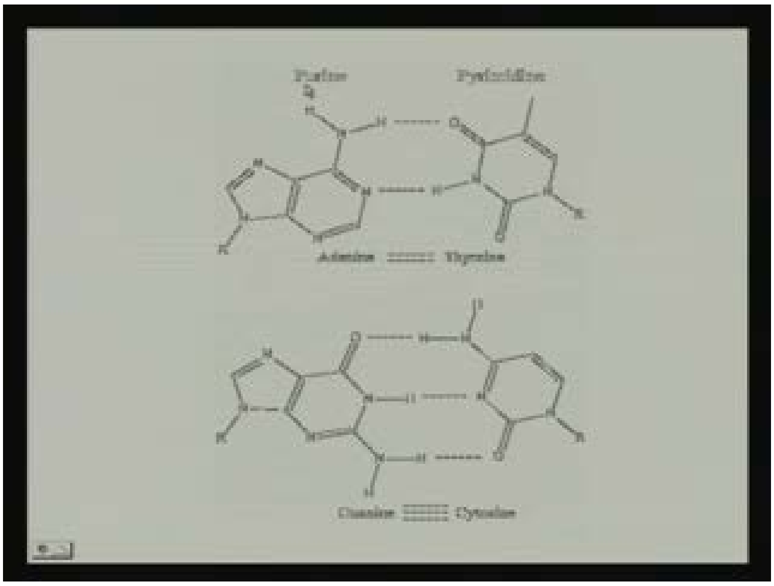
- In DNA there are two strands of nucleotides that wind together in a **double helix**
  - the strands run in opposite directions i.e antiparallel
  - the bases are arranged in step-like pairs
  - the **base pairs** are held together by hydrogen bonding
- The pairing of the bases from the two strands is very specific
- **The complimentary base pairs** are A-T and G-C
  - two hydrogen bonds form between A and T
  - three hydrogen bonds form between G and C
- Each pair consists of a purine and a pyrimidine, so they are the same width, keeping the two strands at equal distances from each other

Now, coming to the secondary structure, so when we are talking about the this primary structure. We are, we have already learned. But, this is a very simple structure which is there and joining the nucleotide together in the secondary structure. It we can find the D N A double helix, that means in D N A there are two strands of nucleotides. That wind together in a double helix form the strands run in opposite direction that is anti parallel to each other .What does it mean? That means if one strand is from 5 prime to 3 prime another strand will be from 3 prime to 5 prime.

So, this opposite link that numbers that 5 prime, if it is starting another end will be 3 prime. And when this two strands are coming, that is that means they are anti parallel to each other .When they come and they wind each other and form this type of structure it is called the helix like structure. And, this is the secondary structure of the D N A molecule. The bases are arranged in a step like manner and the base pairs are held together by hydrogen bonding. The base pairing, the pairing of the basis from the two strands is very specific, the complimentary base pairs are A T and G C. That means one purine is getting linked with another pyrimidine that is the linkage no two purine or two pyrimidines are there, one purine is getting linked with another pyrimidine.

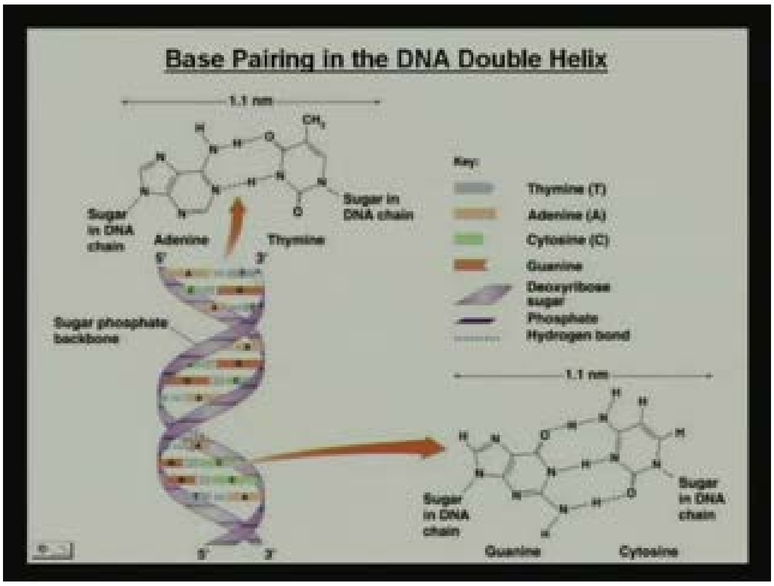
So A is getting bonded with T, G is getting bonded with C, so two hydrogen bonds are formed between A and T. And 3 hydrogen bonds are formed between G and C, each pair consists of purine and pyrimidine. So they are the same with keeping the 2 strands at equal distance from each other.

(Refer Slide Time: 26:47)



So, this is the adenine and this is the thymine, one is purine another pyrimidine. So 2 bonding hydrogen bonding is there, that means A double bond T and G triple bond C. See here G guanine that 1 2 3 hydrogen bonding and cytosine.

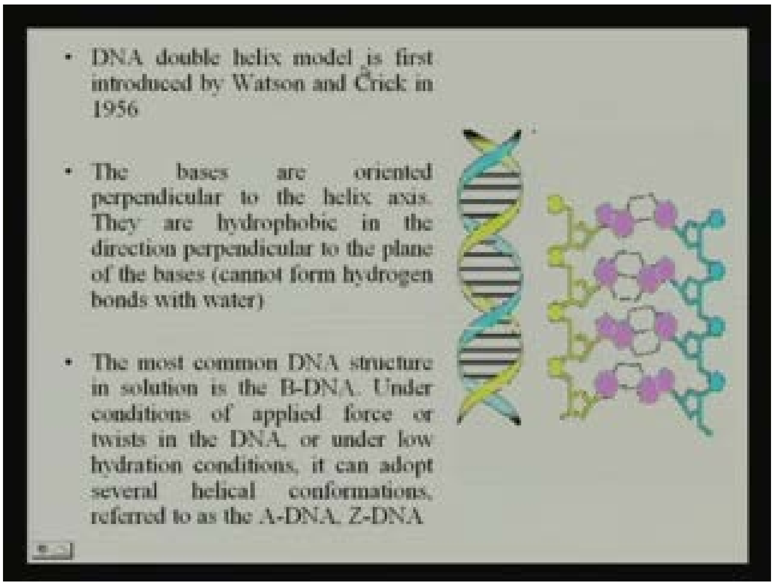
(Refer Slide Time: 27:12)



Base Pairing in the DNA Double Helix

So, these are the helix structure double helical structure of this D N A .A T G C linkages, these pairs are there and which are associated, associating the two strands in a organized manner.

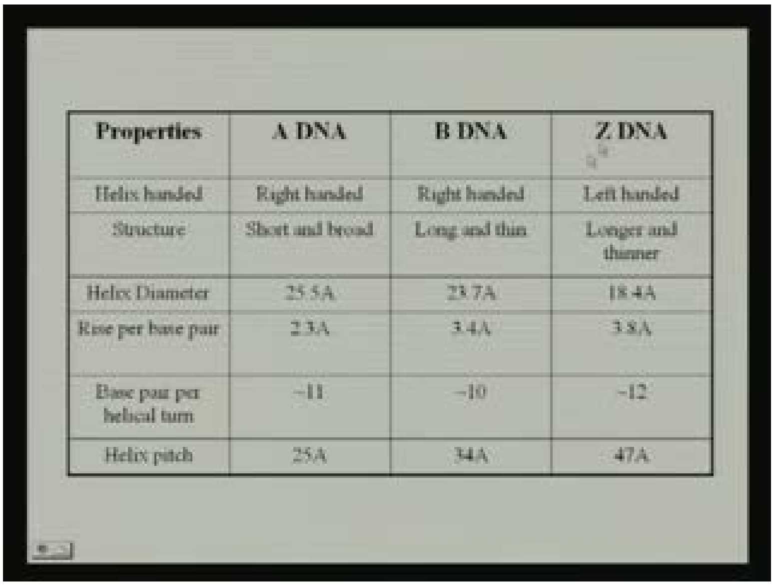
(Refer Slide Time: 27:29)



- DNA double helix model is first introduced by Watson and Crick in 1956
- The bases are oriented perpendicular to the helix axis. They are hydrophobic in the direction perpendicular to the plane of the bases (cannot form hydrogen bonds with water)
- The most common DNA structure in solution is the B-DNA. Under conditions of applied force or twists in the DNA, or under low hydration conditions, it can adopt several helical conformations, referred to as the A-DNA, Z-DNA

D N A double helix model is first introduced by watson and crick in nineteen hundred and fifty six. And they got noble prize for that the basis are oriented perpendicular to the helix axis. They are hydrophobic in the direction perpendicular to the plane of the basis cannot form the hydrogen bong with water. The most common D N A structure in solution is the B form of D N A. Under conditions of applied force or twists in the D N A or under low hydration conditions it can adopt several helical confirmation which are refer to as a form of D N A or z D N A and so on . See here is one nucleotide strand another and you see a T G C G C A T, and that way this helix structure is there. It is formed in between and they are associating the 2 strand together..

(Refer Slide Time: 28:44)

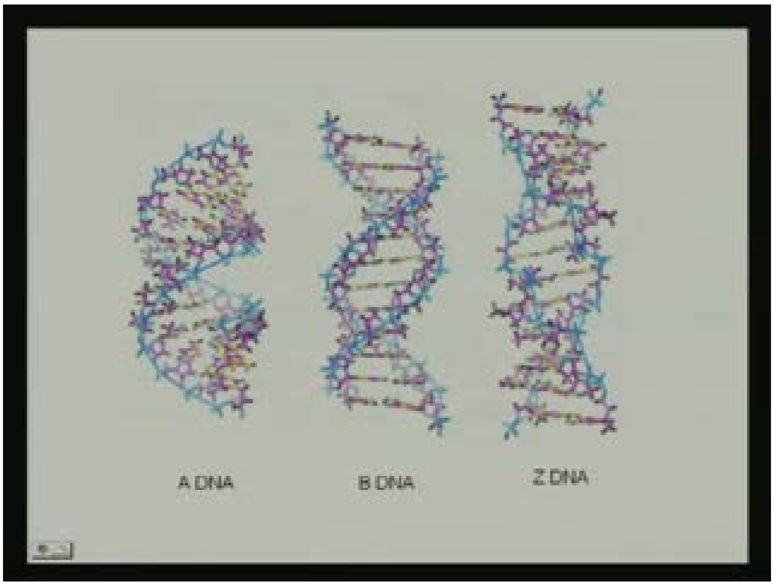


Properties	A DNA	B DNA	Z DNA
Helix handed	Right handed	Right handed	Left handed
Structure	Short and broad	Long and thin	Longer and thinner
Helix Diameter	25.5A	23.7A	18.4A
Rise per base pair	2.3A	3.4A	3.8A
Base pair per helical turn	~11	~10	~12
Helix pitch	25A	34A	47A

While coming to this A form B form and z form. So B form is the normal D N A molecules, and under extreme environmental or extreme conditions. We can get A form of D N A or z form of D N A.

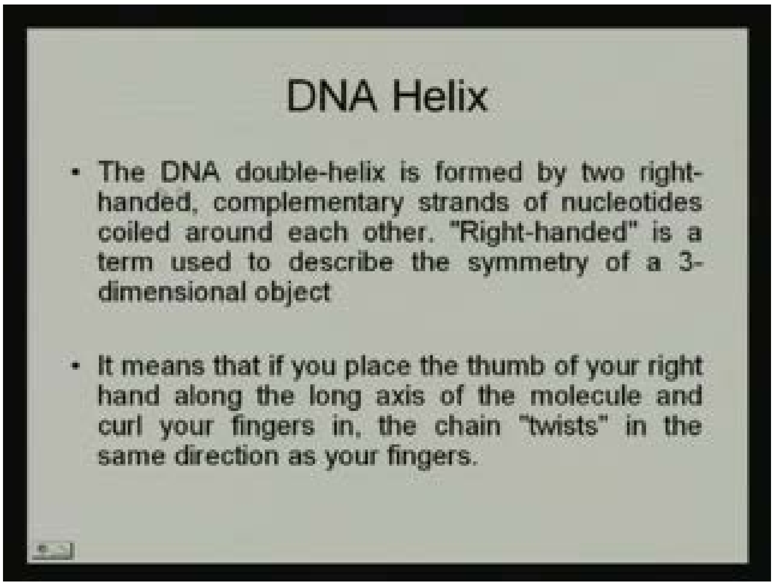
When we are comparing the properties of this type of D N A’s .If we consider the B form as the standard, we can find that this A form. It is the right handed B form is right handed helix, z form is left handed helix. And there are very many differences between this base pair. This helix diameter rise per base pairs, base pair per helical turn, helix pitch etcetera at different in each form of D N A.

(Refer Slide Time: 29:43)



Now, if we see at a glance, we can find that A form of D N A is very thick and short. D if we are considering as normal, then z form is thin and long. So these are some of this D N A helix.

(Refer Slide Time: 29:43)



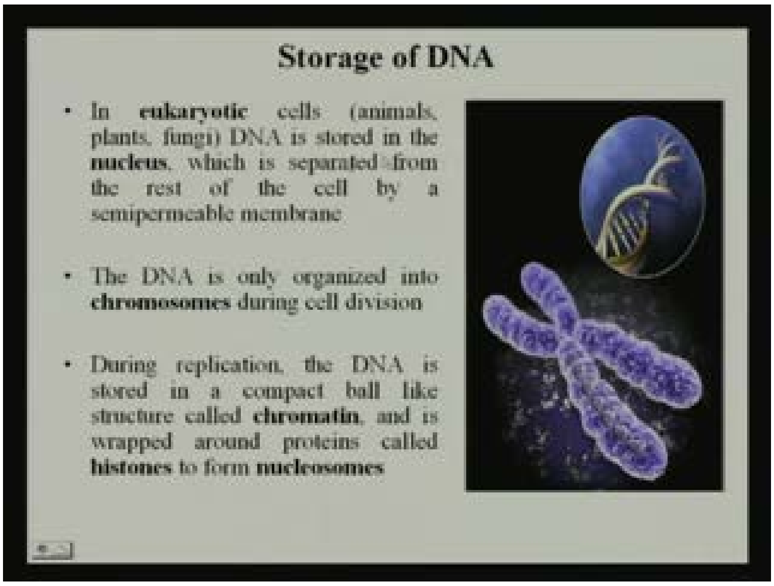
DNA Helix

- The DNA double-helix is formed by two right-handed, complementary strands of nucleotides coiled around each other. "Right-handed" is a term used to describe the symmetry of a 3-dimensional object
- It means that if you place the thumb of your right hand along the long axis of the molecule and curl your fingers in, the chain "twists" in the same direction as your fingers.

The D N A double helix is formed by 2 right handed complimentary strand of nucleotide coil around each other. Right handed is the turn used to describe the symmetry of the 3 dimensional objects. It means that if you place the thumb of your right hand along the long axis of the molecule and curl your fingers in the chain twists in the same direction as your finger.

Now, if suppose this is the double helical structure, and now if we wind this double helix structure to the thumb. And if it is forming this right handed helix, it will be wind like this and this is called the right handed helix of D N A. And in case of left handed, this helical, double helical structure is winds and it forms and in this way it winds up with the central axis, and this is the left handed helix of D N A.

(Refer Slide Time: 30:37)

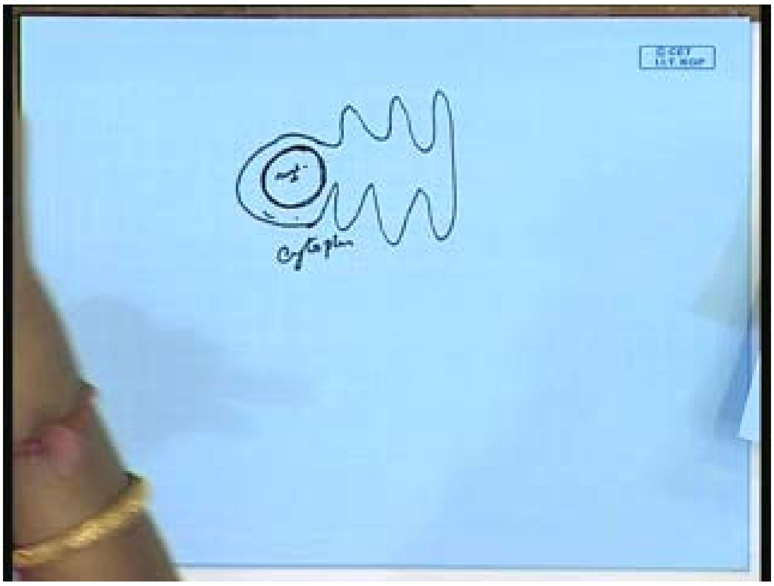


Storage of DNA

- In **eukaryotic** cells (animals, plants, fungi) DNA is stored in the **nucleus**, which is separated from the rest of the cell by a semipermeable membrane
- The DNA is only organized into **chromosomes** during cell division
- During replication, the DNA is stored in a compact ball like structure called **chromatin**, and is wrapped around proteins called **histones** to form **nucleosomes**

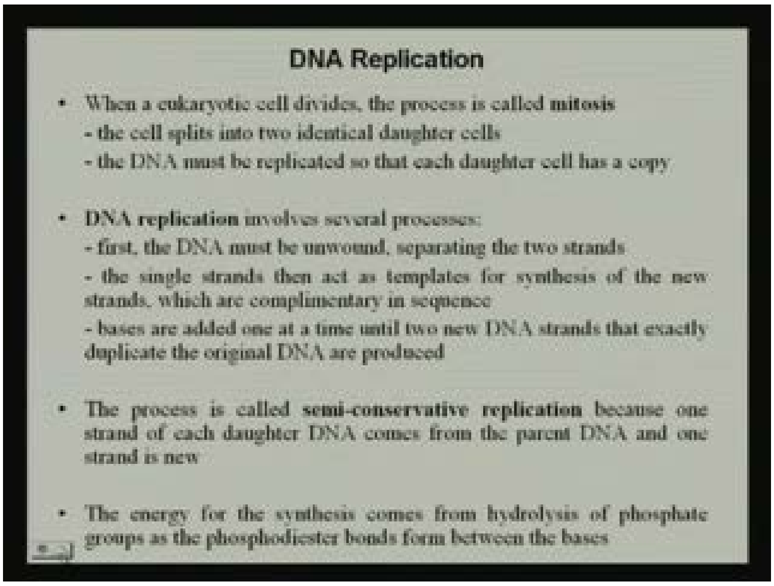
So, if we see this other properties the storage of D N A in eukaryotic cells mainly the animal, plant and fungi .D N A is stored in the nucleus which is separated from the rest of the cell by the cell membrane. And within that the cell sap is called nucleoplasm that is that I have already discussed that thing .That in the inside the nucleus the in case of eukaryotes that nucleus has got its own nuclear membrane and that is separating this nucleus from the other organelles.

(Refer Slide Time: 32:05)



And nucleus has got the by layer that nucleus has got its own , this is one ,this nuclear envelope which is separating the nucleoplasm from this .Other the cytoplasm field and here inside this the nucleic acid or the chromosomes are there . D N A is organized into chromosome during cell division; during replication the D N A is stored in a compact B like structure called the chromatin. And,is wrapped around the protein which is the histone protein to form nucleosomes. So these are some of these things.

(Refer Slide Time: 33:06)



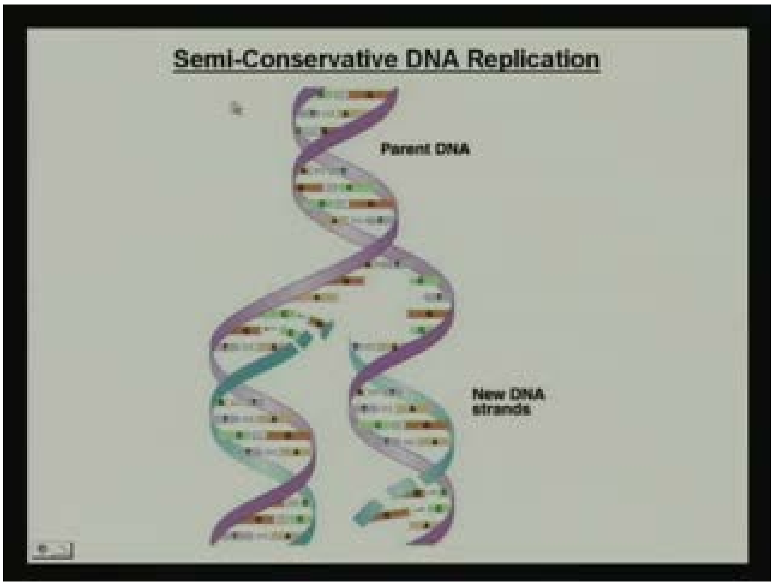
DNA Replication

- When a eukaryotic cell divides, the process is called **mitosis**
  - the cell splits into two identical daughter cells
  - the DNA must be replicated so that each daughter cell has a copy
- **DNA replication** involves several processes:
  - first, the DNA must be unwound, separating the two strands
  - the single strands then act as templates for synthesis of the new strands, which are complimentary in sequence
  - bases are added one at a time until two new DNA strands that exactly duplicate the original DNA are produced
- The process is called **semi-conservative replication** because one strand of each daughter DNA comes from the parent DNA and one strand is new
- The energy for the synthesis comes from hydrolysis of phosphate groups as the phosphodiester bonds form between the bases

So, inside this nucleus the replication of D N A is taking place. When one eukaryotic cell divides, the process is that the process, they adopt is the mitotic cell division, The cell in this particular cell division, the cells split into 2 identical daughter cells, the D N A must be replicated. So that each daughter cell has the same copies. Number of this nuclide this chromosome D N A replication involves in several processes.

First, is the D N A must be unwound and separating the two strands. The single strands then act as a template for synthesis of the new strand which are complimentary in sequence. The bases are added one at a time until the two new strands are formed .That exactly duplicate the original D N A which are produced .The process is called semi conservative type of D N A replication, because one strand of each daughters D N A comes from the parent D N A and one strand is synthesize which is new strand.During this process the energy which is required comes from the hydrolysis of the phosphate group .As the phosphodiester bond form between the basis, as we have already learned that thing.

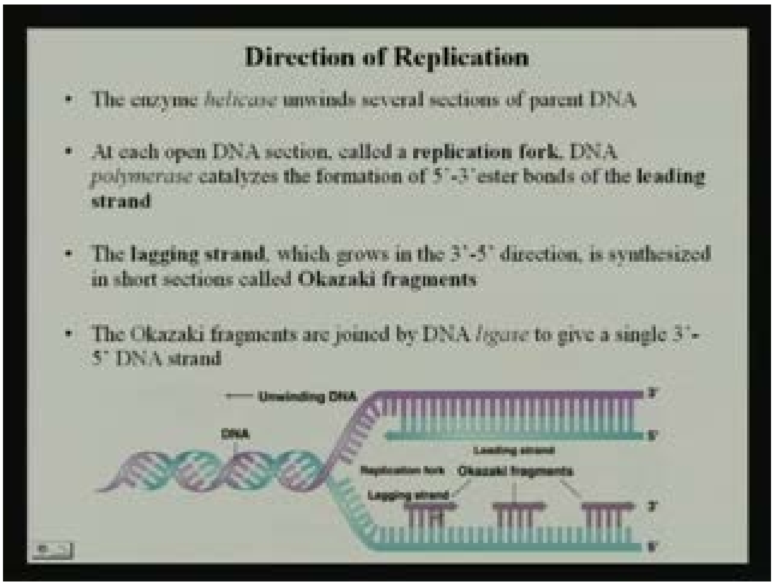
(Refer Slide Time: 34:56)



Semi-Conservative DNA Replication

Now, if we are talking about this semi conservative D N A replication. Then this is the parents strand, and first what it is doing? It gets unwind now one is the 3 prime to 5 prime another it is 5 prime to 3 prime. So, anti parallel to each other, so when these two strands are getting separated, then new strand, this own strand is the template strand for synthesis of the new strand. And this way, this new strands are synthesized and when this new strands are synthesized .It has got one old the template strand, parental strand and new synthesize strand.

(Refer Slide Time: 35:50)



Direction of Replication

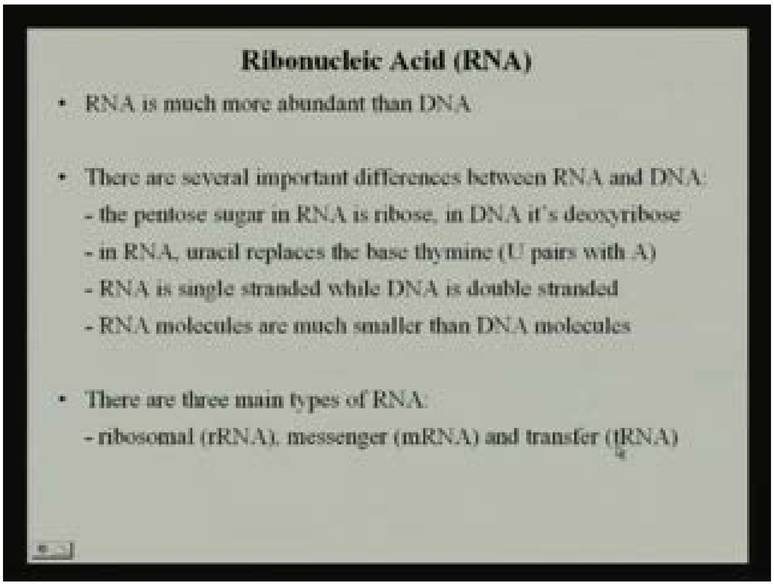
- The enzyme *helicase* unwinds several sections of parent DNA
- At each open DNA section, called a **replication fork**, DNA *polymerase* catalyzes the formation of 5'-3' ester bonds of the **leading strand**
- The **lagging strand**, which grows in the 3'-5' direction, is synthesized in short sections called **Okazaki fragments**
- The Okazaki fragments are joined by DNA *ligase* to give a single 3'-5' DNA strand

And these are called the semi conservative type of D N A replication. The enzyme helicase unwinds the several section of the parent D N A at each open D N A section called a replication fork.

So, here this is the open up form that is the replication form of D N A. Polymerase catalyzes the formation of 5 prime to 3 prime ester bond of the leading strand. So here this is the parental strand got opened so one end is the 3 prime end another is the 5 prime end. This is the one parental strand, this is another parental strand.

Now, there is a leading strand and another is the lagging strand. So when 3 prime to this 3 prime end is there so 5 prime to 3 prime, this leading strand is start synthesizing and they form the complimentary base pair of this particular new strand. Where as in the lagging strand which grow from 3 prime to 5 prime direction is synthesized a small short section of this nucleotide base pair. And that is called the okazaki fragment ,this okazaki fragment with the help of inside D N A ligase they join and to form a single 3 prime to 5 prime strand .And in this way the replication process is taking place in D N A.

(Refer Slide Time: 37:30)



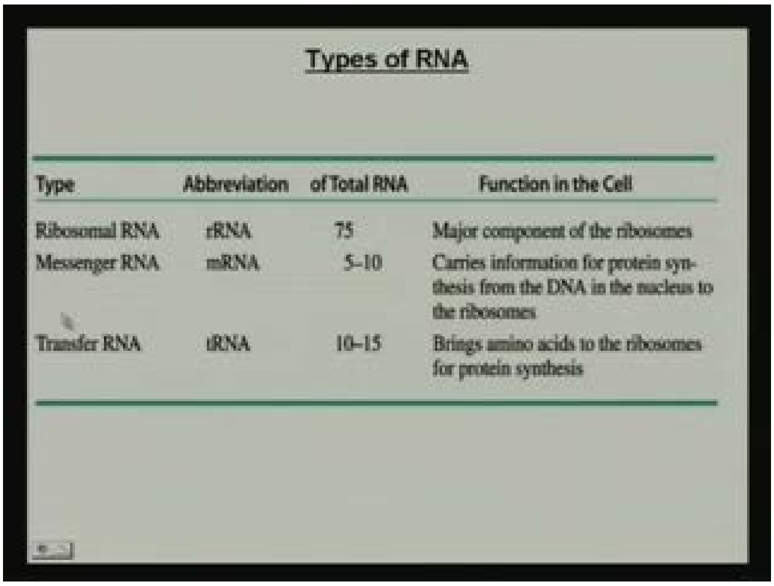
Ribonucleic Acid (RNA)

- RNA is much more abundant than DNA
- There are several important differences between RNA and DNA:
  - the pentose sugar in RNA is ribose, in DNA it's deoxyribose
  - in RNA, uracil replaces the base thymine (U pairs with A)
  - RNA is single stranded while DNA is double stranded
  - RNA molecules are much smaller than DNA molecules
- There are three main types of RNA:
  - ribosomal (rRNA), messenger (mRNA) and transfer (tRNA)

That means the replication means D N A to D N A .So from one parental strain the template is that is used as the template and new synthesized strand is taking place. Coming to this ribonucleic acid or the R N A , R N A is the most abundant than D N A. There are several important differences between R N A and D N A. The pentose sugar in R N A is ribose and in case of D N A is deoxyribose. In R N A uracil replaces the base thymine that means U pair with a R N A is single stranded .While D N A is the double stranded R N A molecules are much smaller than the D N A molecules.

If we categorize this R N A ,then we can differentiate this R N A as ribosomal R N A or r R N A ,messenger R N A or m R N A and transfer R N A or the t R N A.

(Refer Slide Time: 38:50)

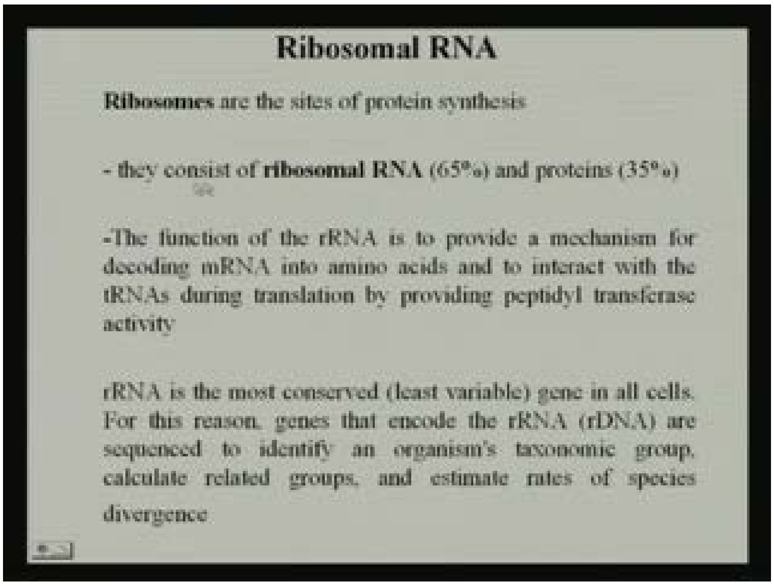


Types of RNA

Type	Abbreviation	of Total RNA	Function in the Cell
Ribosomal RNA	rRNA	75	Major component of the ribosomes
Messenger RNA	mRNA	5–10	Carries information for protein synthesis from the DNA in the nucleus to the ribosomes
Transfer RNA	tRNA	10–15	Brings amino acids to the ribosomes for protein synthesis

Now, if we see in the R N A molecules the different types, different compositional variation of these 3 R N A's. Ribosomal R N A or r R N A of the total R N A present, it constitute 75 percent .Where as m R N A is the 5 to 10 percent and t R N A is 10 to 15 percent of the total R N A which is present. If we see the functions of these particular R N A in the cell inside the cell, then we can find that r R N A is the major component of the ribosome. M R N A carries the information of protein synthesis from the D N A in the nucleus to the ribosome .And t R N A brings amino acid to the ribosome for protein synthesis.

(Refer Slide Time: 39:51)



Ribosomal RNA

Ribosomes are the sites of protein synthesis

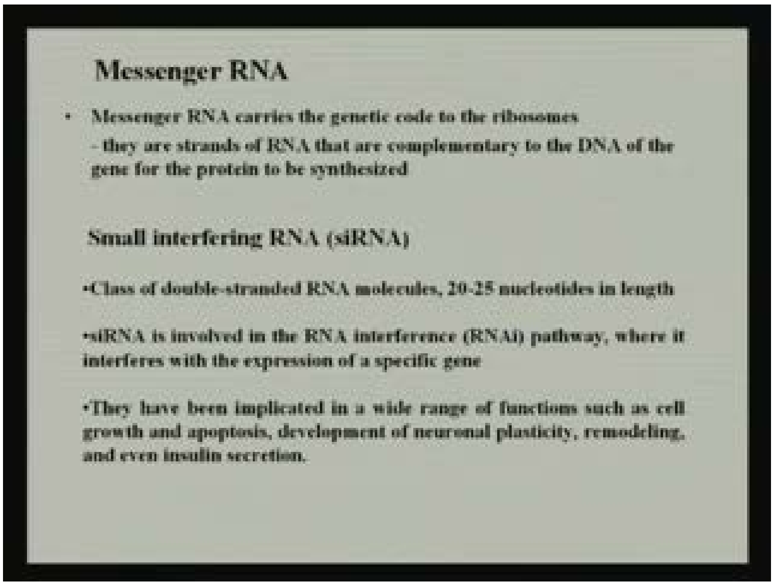
- they consist of **ribosomal RNA** (65%) and proteins (35%)

-The function of the rRNA is to provide a mechanism for decoding mRNA into amino acids and to interact with the tRNAs during translation by providing peptidyl transferase activity

rRNA is the most conserved (least variable) gene in all cells. For this reason, genes that encode the rRNA (rDNA) are sequenced to identify an organism's taxonomic group, calculate related groups, and estimate rates of species divergence

So,this way this this R N A different. R N A's are doing a very important role in cell system. Ribosomal R N A when we are talking about this, is the ribosomes are the side for the protein synthesis .They consists of ribose, they consists of ribosomal R N A 65 percent and protein 35 percent .The function of r R N A is to provide a mechanism for decoding m R N A into amino acids and it interact with the t R N A during translation by providing peptidyl transferase activity. R R N A is the most conserved and that is the reason why now a days we rely upon the r R N A data for identification of any unknown species..

(Refer Slide Time: 40:50)



Messenger RNA

- Messenger RNA carries the genetic code to the ribosomes
  - they are strands of RNA that are complementary to the DNA of the gene for the protein to be synthesized

Small interfering RNA (siRNA)

•Class of double-stranded RNA molecules, 20-25 nucleotides in length

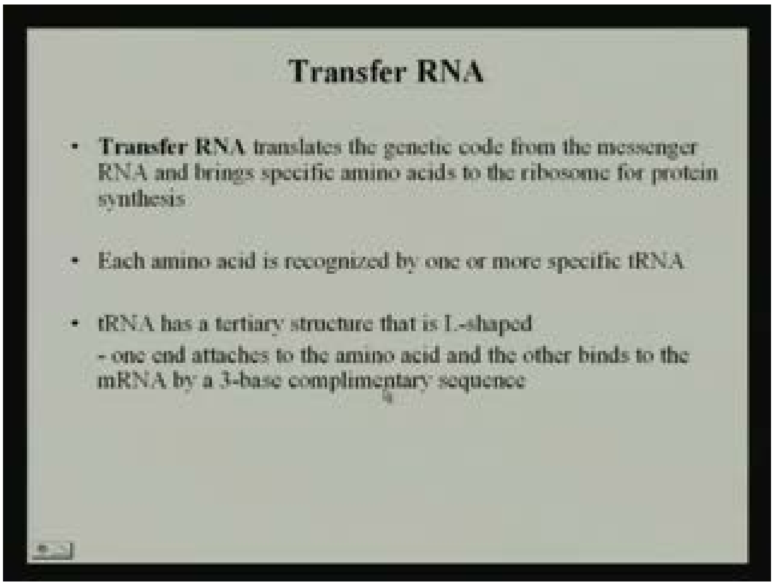
•siRNA is involved in the RNA interference (RNAi) pathway, where it interferes with the expression of a specific gene

•They have been implicated in a wide range of functions such as cell growth and apoptosis, development of neuronal plasticity, remodeling, and even insulin secretion.

So, this is the most important part of r R N A. Messenger R N A is the m R N A carries the genetic code to the ribosome .They are strands of R N A that are complimentary to the D N A of the gene ,for the protein to be synthesized.

There are another type of R N A called small interfering R N A or it is called s i R N A. There are the classes of double stranded R N A molecules, 20 to 25 nucleotide despaired in length. S i R N A is involved in the R N A interference, that is R N A i activities path way .Where it interfere with the expression of a specific gene, they have implicated in the wide range of functions such as cell growth or apoptosis. That is programmed cell death, development of neuronal plasticity, remodeling and even insulin secretion.

(Refer Slide Time: 42:02)

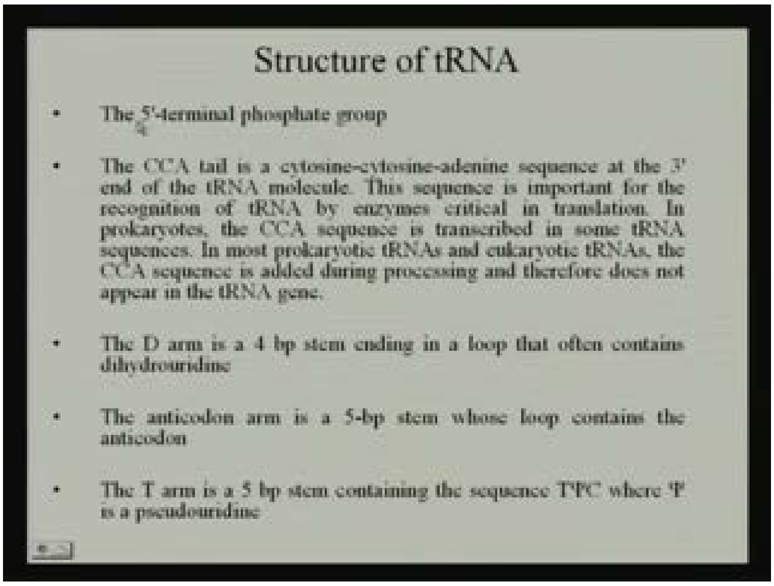


Transfer RNA

- **Transfer RNA** translates the genetic code from the messenger RNA and brings specific amino acids to the ribosome for protein synthesis
- Each amino acid is recognized by one or more specific tRNA
- tRNA has a tertiary structure that is L-shaped
  - one end attaches to the amino acid and the other binds to the mRNA by a 3-base complimentary sequence

Now, coming to the t R N A, transfer R N A translate the genetic code from the messenger R N A and bring specific amino acid to the ribosome for protein synthesis. Each amino acid is recognized by one or more specific t R N A. T R N A has a tertiary structure that is L shaped ,one end is attached to the amino acid and the other bind to the messenger R N A by the 3 base complimentary sequence.

(Refer Slide Time: 42:45)



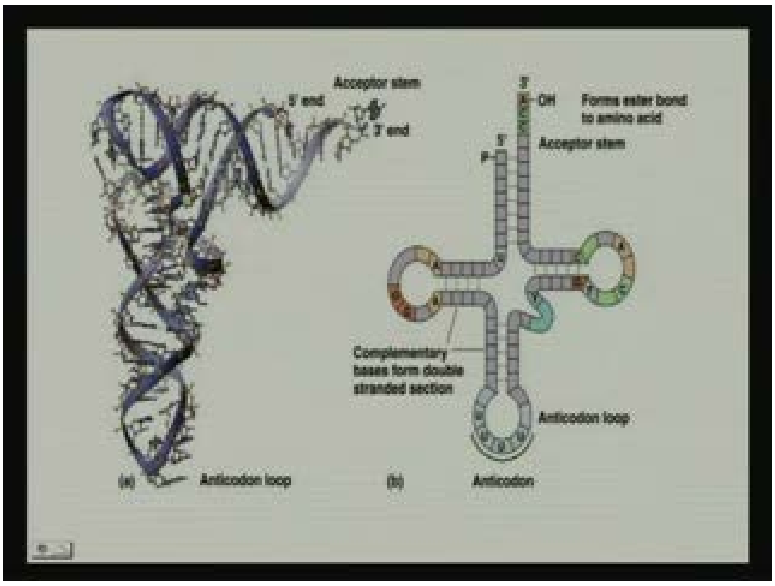
Structure of tRNA

- The 5'-terminal phosphate group
- The CCA tail is a cytosine-cytosine-adenine sequence at the 3' end of the tRNA molecule. This sequence is important for the recognition of tRNA by enzymes critical in translation. In prokaryotes, the CCA sequence is transcribed in some tRNA sequences. In most prokaryotic tRNAs and eukaryotic tRNAs, the CCA sequence is added during processing and therefore does not appear in the tRNA gene.
- The D arm is a 4 bp stem ending in a loop that often contains dihydrouridine
- The anticodon arm is a 5-bp stem whose loop contains the anticodon
- The T arm is a 5 bp stem containing the sequence TΨC where Ψ is a pseudouridine

If, we see the structure of t R N A ,it has got a 5 prime terminal phosphate group .That this particular t R N A has a C C A tail that is cytosine cytosine adenine sequence at the 3 prime end of the t R N A molecule .This sequence is important for the recognition of t R N A by enzymes critical in translation .In prokaryotes the C C A sequence is transcribed in some t R N A sequence, in most prokaryotic t R N A and eukaryotic t R N A’s, the C C A sequence is added during processing .And therefore, does not appear in the t R N A gene.

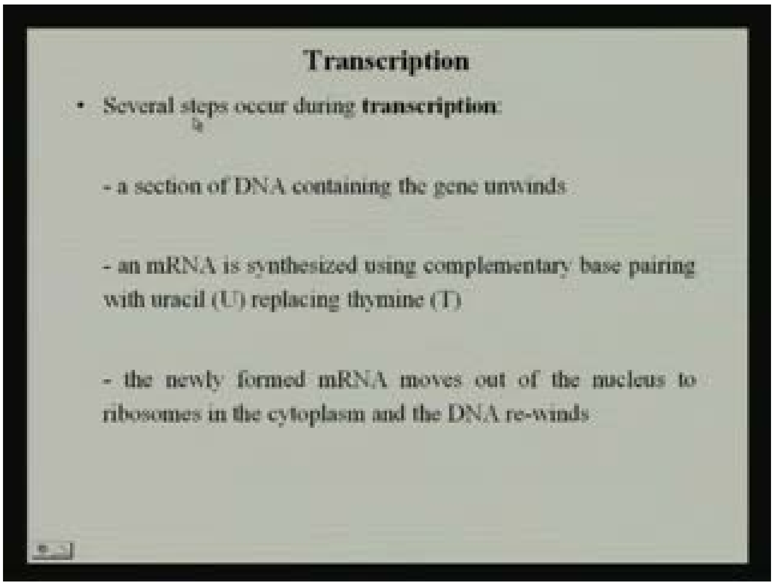
This t R N A has a d arm that is four base pair stem ending in a loop that often contain dihydrouridine. It has got also anticodon arm that is 5 base pair stem, whose loop contains the anticodon .And it has got also a t arm, that is the 5 base pair stem containing the sequence that is t (( )) C where psi is the pseudouridine.

(Refer Slide Time: 44:23)



So, as I have told you already, that it is the t R N A which is L shape .See this is the L shaped t R N A .If we see this is the loop, where this anticodon loop is there. And if we see the structure, that see in the 3 prime end of this particular hydroxyl group, one C C A group is there. It forms the ester bond to the amino acid and this is the acceptor stem, and it has also got a 5 prime end .As I have already discussed that I have also told you that it has got that loops, that one is the d arm another is the t arm .And another is the anticodon, and this anticodon arm is the 5 base pair stem..

(Refer Slide Time: 45:50)



Transcription

- Several steps occur during **transcription**:

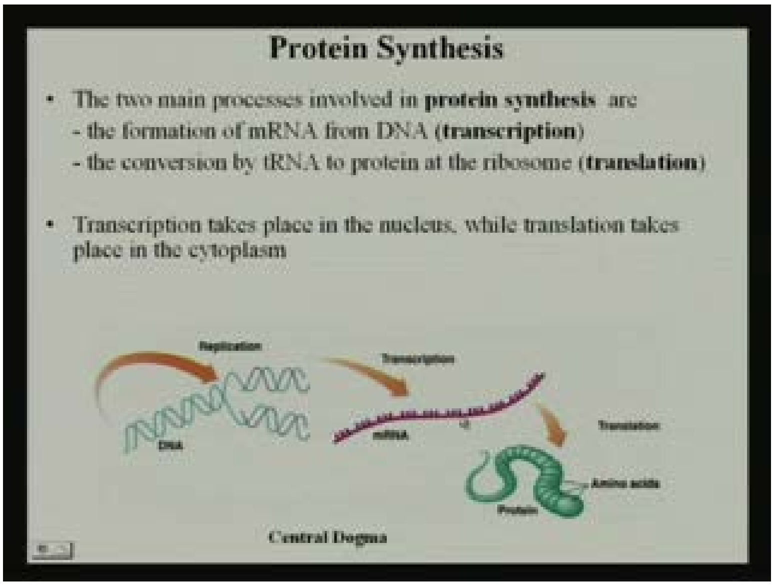
  - a section of DNA containing the gene unwinds

  - an mRNA is synthesized using complementary base pairing with uracil (U) replacing thymine (T)

  - the newly formed mRNA moves out of the nucleus to ribosomes in the cytoplasm and the DNA re-winds

And it sends the messenger R N A which is coming with the polysome and forming the nucleotide. And it is sensing the amino acid for this particular that m R N A and this is the importance of this particular t R N A. And it brings the amino acid to that particular sight. Coming to this transcription that several steps occur during transcription, a section of D N A containing the gene got unwind ,the m R N A is synthesized using complimentary base pairing using uracil . That is U replacing the thymine, that is T the newly formed m R N A moves out of this nucleus to the ribosome in the cytoplasm and the D N A rewind.

(Refer Slide Time: 46:30)



Protein Synthesis

- The two main processes involved in **protein synthesis** are
  - the formation of mRNA from DNA (**transcription**)
  - the conversion by tRNA to protein at the ribosome (**translation**)
- Transcription takes place in the nucleus, while translation takes place in the cytoplasm

Central Dogma

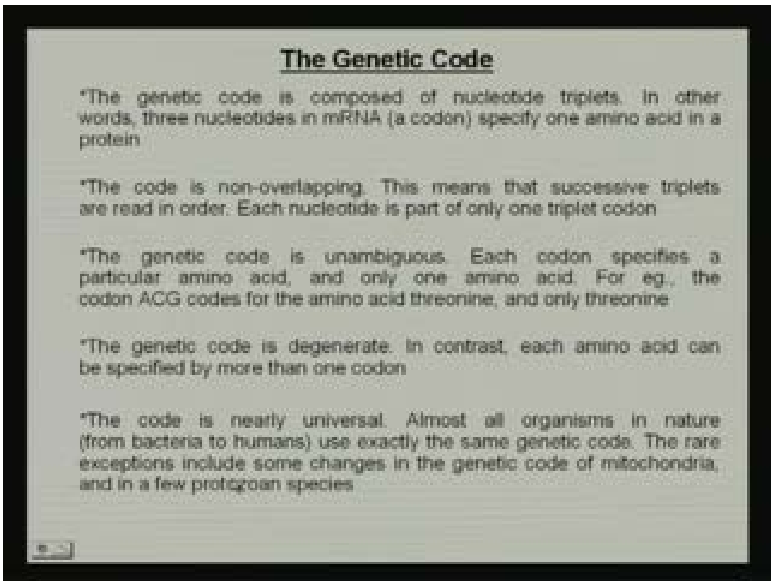
So, if we see I have already discussed that the D N A replication is taking place in the nucleus..

(Refer Slide Time: 46:40)

D N A replication is taking place, transcription when it takes place, D N A form the m R N A. When m R N A is formed m R N A comes out of this nucleus, to this outside of the cell and it is coming to the cytoplasmic field of the cell. And these two main processes involved in protein synthesis are the formation of m R N A. From this D N A that is the transcription and conversion of t R N A to proteins at the ribosome that is the translation transcription takes place in nucleus, while translation takes place in the cytoplasm.

So, as soon as m R N A is formed, m R N A is coming out from this nucleus to the cytoplasm. Where this translation is taking place? So, here if we see the transcription central dogma of the cell ,see here replication is taking place .D N A to D N A is the replication ,when D N A is undergoing transcription it forms D N A, forms this m R N A. When m R N A is formed, this m R N A comes out from this nucleus to the cytoplasmic field. Where actual protein synthesis is going on and these amino acids are producing this protein from the translation process ,which is going on in the cell with the help of this nucleic acid.

(Refer Slide Time: 48:24)



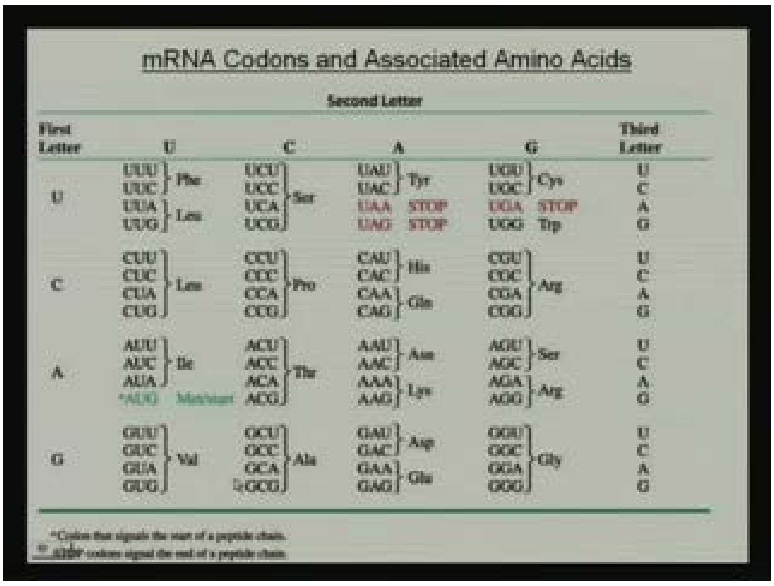
The Genetic Code

- The genetic code is composed of nucleotide triplets. In other words, three nucleotides in mRNA (a codon) specify one amino acid in a protein
- The code is non-overlapping. This means that successive triplets are read in order. Each nucleotide is part of only one triplet codon
- The genetic code is unambiguous. Each codon specifies a particular amino acid, and only one amino acid. For eg., the codon ACG codes for the amino acid threonine, and only threonine
- The genetic code is degenerate. In contrast, each amino acid can be specified by more than one codon
- The code is nearly universal. Almost all organisms in nature (from bacteria to humans) use exactly the same genetic code. The rare exceptions include some changes in the genetic code of mitochondria, and in a few protozoan species

Now, coming to this genetic code, the genetic code is composed of nucleotide. Triplets in some, in other word 3 nucleotides of m R N A ,that is specify one amino acid in a particular protein .So, what we can tell that 3 nucleotide. If a codon , which specify what one particular amino acid ,one codon is specify with one particular amino acid. The code is non-overlapping, that means there is no over lapping of this particular code, this means that successive triplets are read,in order each nucleotide is part of only one triplet codon..

The genetic code is unambiguous, each codon specifies a particular amino acid and only one amino acid for example, the codon a C G codes for the amino acid threonine and only threonine. The genetic code is degenerate in contrast, each amino acid can be specified by more than one codon .Each code is nearly universal in nature, almost all organisms in nature from bacteria to human being are exactly the same genetic use, exactly the same genetic code. The real exceptions include some changes in the genetic code of mitochondria and in a few protozoan species.

(Refer Slide Time: 50:35)



mRNA Codons and Associated Amino Acids

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

*Codon that signals the start of a peptide chain.

** STOP codons signal the end of a peptide chain.

Now, coming to this m R N A codon and associated amino acids. So I have already told that this genetic codes are the triplet codes, so if we are taking this first letter, second letter and third letter, this that nitrogenous base pairs, that see here, how these codes are formed? That C U C A G this is first letter U C A G and this is the third letter U C A G.

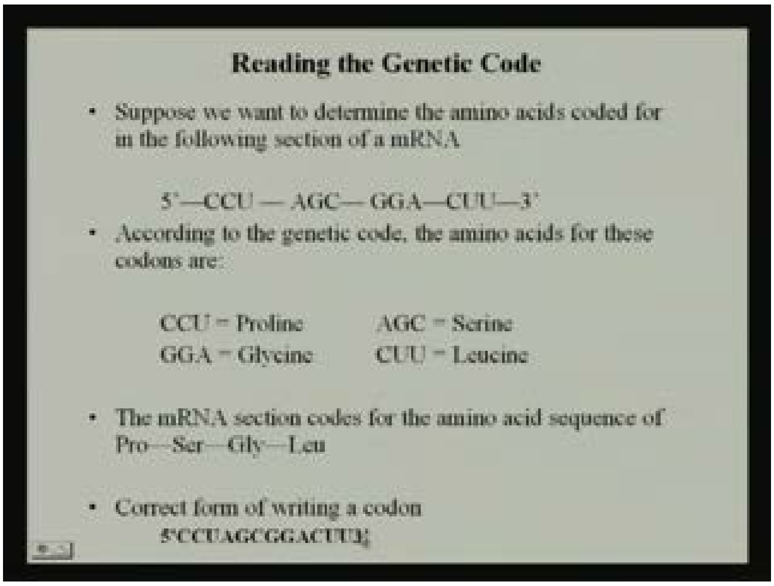
Now, if we form this genetic code, then 3 letter triplet codes are there U U U , U U C ,U U A, U U G see U U U, U U C, U U U, U U A and U U G .So, these are the codes triplet codes. Now, in reality, we have four base pairs and 3 nucleotides; form this one code, so in nature we have sixty four such nucleotides. These codes which are coding this amino acid. But, in reality, what is happening in reality? We have only twenty standard amino acid. I have already two and they form this permutation and combination with their permutation and combination, these proteins are formed.

So, what is happening in reality? If this 64 codons are there and twenty amino acids are there. And if each codon are responsible for coding , only one amino acid, then what will happen? There will be 44 such codons that this codons ,for which codons will be there .So during protein synthesis, when it will be coming with any permutation and combination and if this forty four codons. If there is no amino acids are there, so any point of time when such codons will be coming, no amino acids will be there for such type of codon .And any time this protein that chain synthesis will get terminated ,but, in reality it is not happening like that. That means, what is there in reality? For each codon there are certain amino acids which are there.

Say for example, for G G U , G G C, G G A, G G G ,for this four codon only one amino acids are there, the which is coding glycine A G U ,A G C, A G A ,A G G .See A G U A G C is serine. A G A A G G is arginine, and in this way for this natural this 64 codon there are certain amino acids which are coded. And that is the reason why it is called it is the the, it is degenerating. So when but, there are some exceptions like a U G is the start codon which codes for the methionine and U A A ,U A G, U G A are called the stop codon. And only one code which only codes the trypsin is the U G G , this is the exception of the genetic code.

Now, if we see another example, say we can let us take any example C G U, C G C, C G A, C G G. Now, here if we see this four codons are coding one amino acid that is arginine. If we see that see this first two codons are same for this particular amino acid, but, the third codon is varying. In case of glycine also you see first 2 codons are constant, same for the, whereas, the third amino acid is varying. That means the third codon is not that important, that third nucleotide base pair is not that important for this particular recognition of this amino acid? But, first two are very important and that is the reason that first two are constant for identify the particular amino acid, and that is called the wobble hypothesis of genetic code.

(Refer Slide Time: 55:53)



Reading the Genetic Code

- Suppose we want to determine the amino acids coded for in the following section of a mRNA

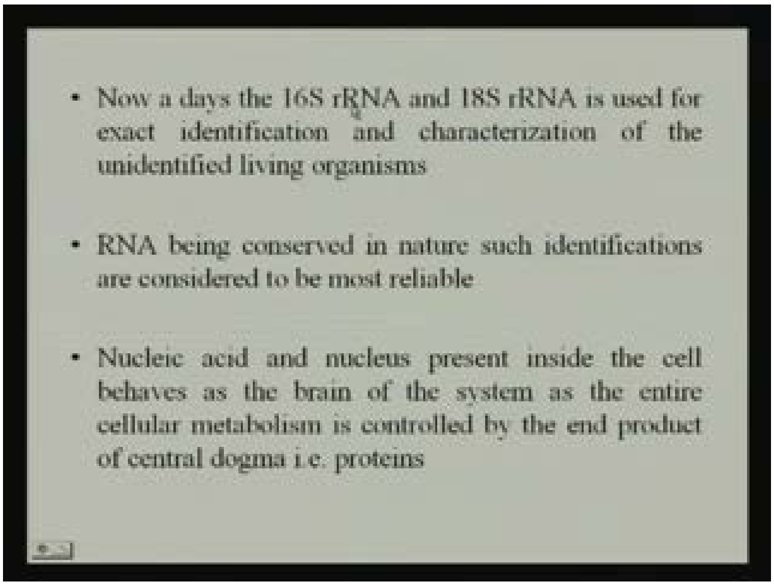
  5'—CCU — AGC— GGA—CUU—3'
- According to the genetic code, the amino acids for these codons are:

  CCU = Proline AGC = Serine
  GGA = Glycine CUU = Leucine
- The mRNA section codes for the amino acid sequence of Pro—Ser—Gly—Leu
- Correct form of writing a codon
  5'CCUAGCGGACUU3'

Now, if we come to this reading of this genetic code ,suppose we want to determine the amino acids coded for this following section that is 5 prime C C U dash A G C dash G G A dash C C U dash 3 prime .Where we know proline is C C U .G G A is glycine .A G C is serine and C U U is leucine.

Now, if we want to write this proline, serine ,glycine ,leucine .We can write but, when we are writing the actual codon ,we have to have the 5 prime C C U, A G C ,G G A, C U U and 3 prime. While writing this codon there should not be any punctuation, any comma, any full stop, any dash or any such type of use of punctuation, while writing the genetic code .And this is the correct way of writing the codon.

(Refer Slide Time: 57:02)



- Now a days the 16S rRNA and 18S rRNA is used for exact identification and characterization of the unidentified living organisms
- RNA being conserved in nature such identifications are considered to be most reliable
- Nucleic acid and nucleus present inside the cell behaves as the brain of the system as the entire cellular metabolism is controlled by the end product of central dogma i.e. proteins

Nowadays, 16 s r R N A for prokaryotic and 18 s r R N A is used for identification and characterization of unidentified living organisms. As because, they are conserved in nature, R N A being conserved in nature, such identification are considered to be the most reliable. Nucleic acid and nucleus present inside the cell behaves as the brain of the system, as the entire cellular metabolism is controlled by the end product of the central dogma. That is the protein and with this, today concluding my lecture and I have tried to give you some idea about that, How cell is undergoing division, cell division, growth, replication and reproduction is going on with the help of this nucleic acid this D N A and R N A molecules?

In my next class, I will discuss about another macromolecules, which is playing a significant role in cellular activities, that is the lipid molecules. Thank you very much.