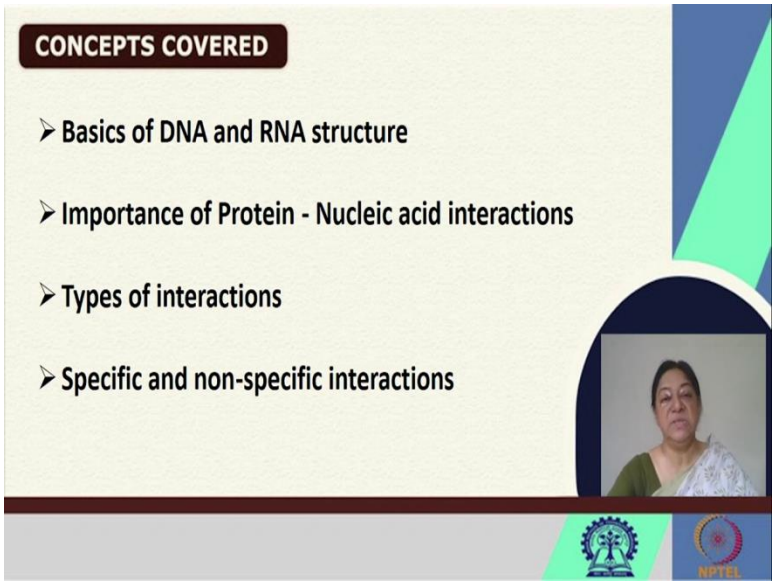
Fundamentals of Protein Chemistry
Prof. Swagata Dasgupta
Department of Chemistry
Indian Institute of Technology, Kharagpur

Module - 10
Protein Macromolecule Interactions I
Lecture - 48
Protein Nucleic Acid Interactions - I

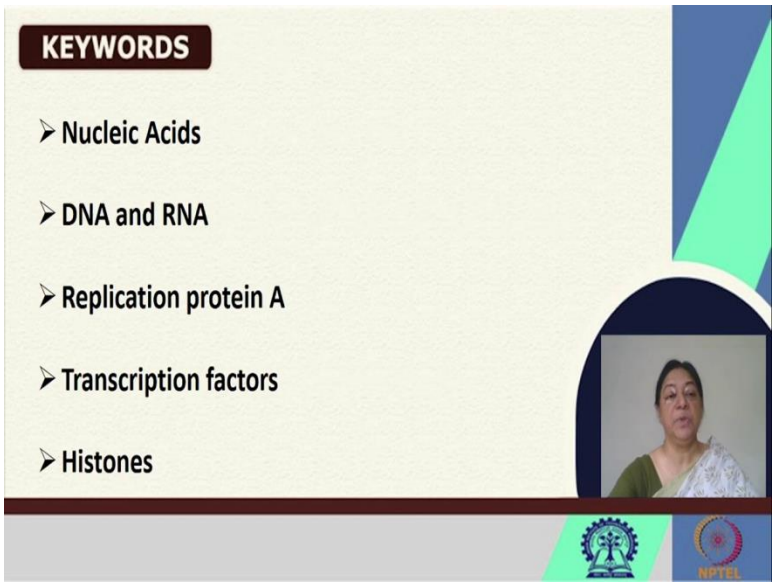
In the next two modules we will be speaking of protein molecule interactions, by looking at several other macromolecules and how they interact with proteins. The first three of these lectures will be on protein nucleic acid interactions; why we need to study them and what their importances are in terms of the mechanism and function of life processes.

(Refer Slide Time: 00:44)

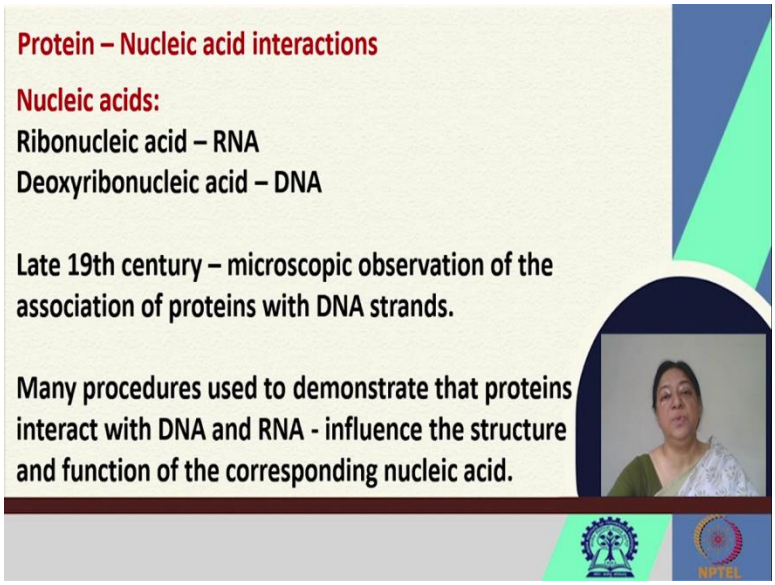


So if we look at this specific set, we will be looking at the basics of DNA and RNA structure, the importance of protein nucleic acid interactions, the types of interactions and what we mean by specific and non-specific interactions.

(Refer Slide Time: 00:59)

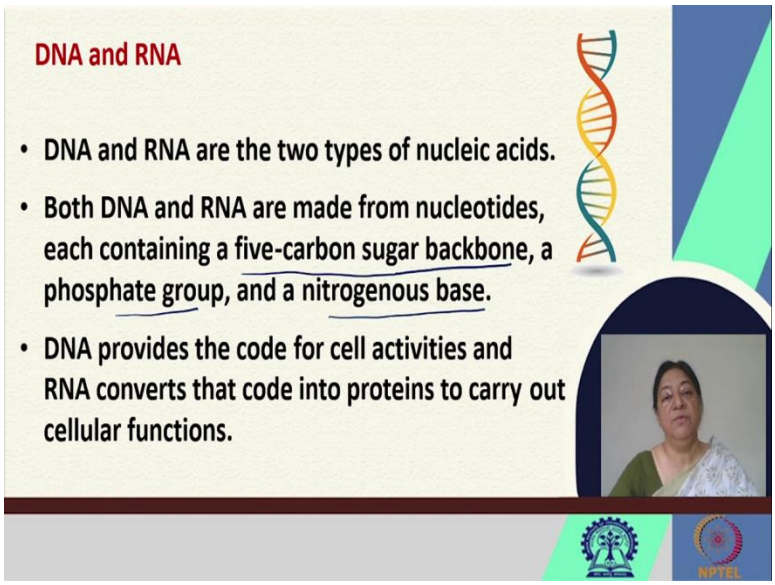


(Refer Slide Time: 01:02)



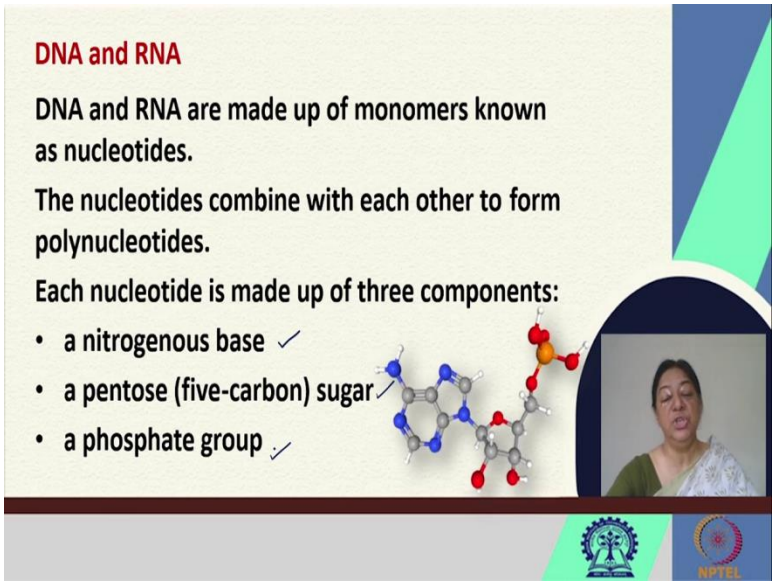
It was discovered late in the 19th century that there were microscopic observations of association of proteins with the DNA strands. If we look at these nucleic acids, we have ribonucleic acid RNA and deoxyribonucleic acid DNA. There are many procedures now, that are used to demonstrate that proteins do interact with DNA and RNA and during the interaction, they influence the structure and the function of the corresponding nucleic acid that they are associated with.

(Refer Slide Time: 01:44)



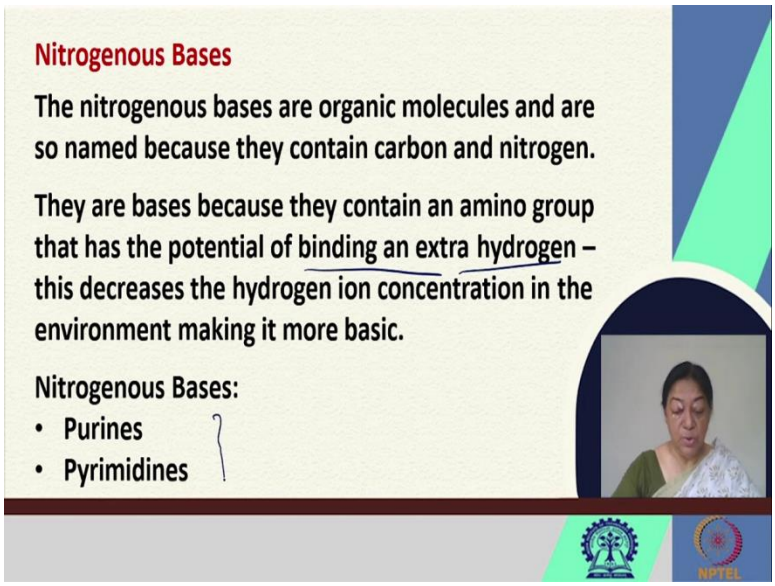
We will look briefly at the structures of DNA and RNA and see how they are important in the interaction with proteins. These are two types of nucleic acids; both RNA and DNA are made from nucleotides. The nucleotides contain the five carbon sugar backbone, a phosphate group and a nitrogenous base. We will see what these mean in a moment. DNA provides the code for all cell activities and RNA converts that code into proteins, to carry out cellular functions in our central dogma of biology, from DNA to RNA to protein.

(Refer Slide Time: 02:32)



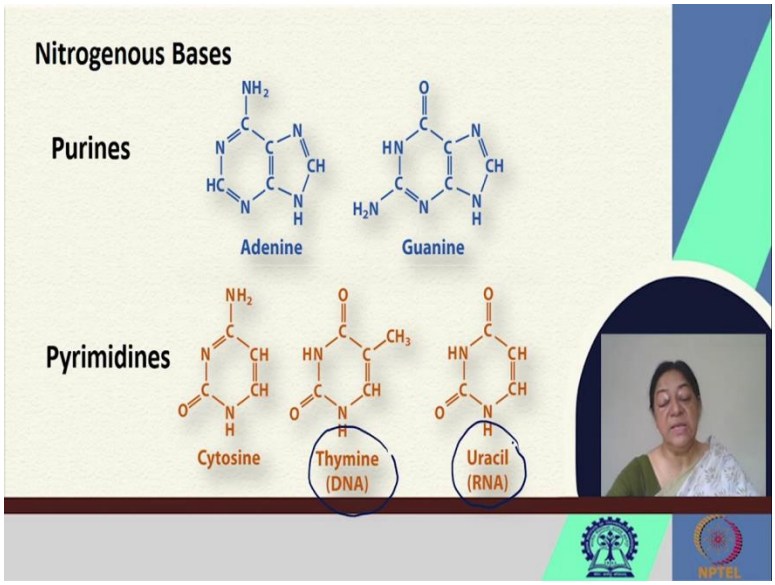
When we look at the DNA and RNA structures, the nucleotides as we saw, were composed of a nitrogenous base, a pentose (five carbon sugar) and a phosphate group. Now, the nucleotides themselves combine with each other in what are called phosphodiester bonds, to form polynucleotides in longer chains of the mononucleotide.

(Refer Slide Time: 03:02)



The nitrogenous bases that are involved in the specific structures of DNA and RNA, are of specific types. They are organic molecules that are named because they have carbon and nitrogen and they are bases because they have the presence of an amino group, that could have the binding possibility of hydrogen bonding, which is extremely important. Association of the nucleic acid strands in DNA, is due to this hydrogen bonding. The nitrogenous bases that we have are the purines and the pyrimidines.

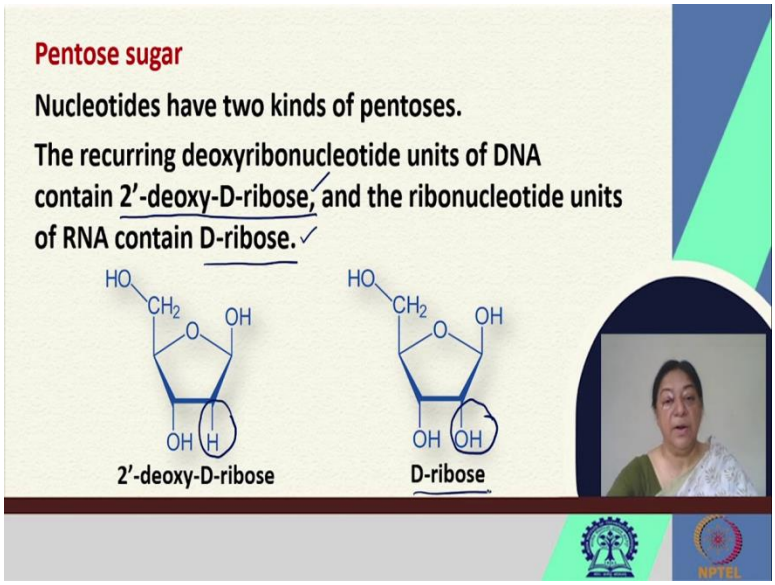
(Refer Slide Time: 03:53)



The purines are adenine and guanine and the pyrimidines are cytosine, thymine and uracil. Thymine is specifically present in DNA and uracil is present in RNA. So, each of these nucleic acids have their own characteristic properties. They can be combined to form the polynucleotides

in different fashions, in the sense that they could interact in forming the specific polynucleotide chain through the phosphodiester bond.

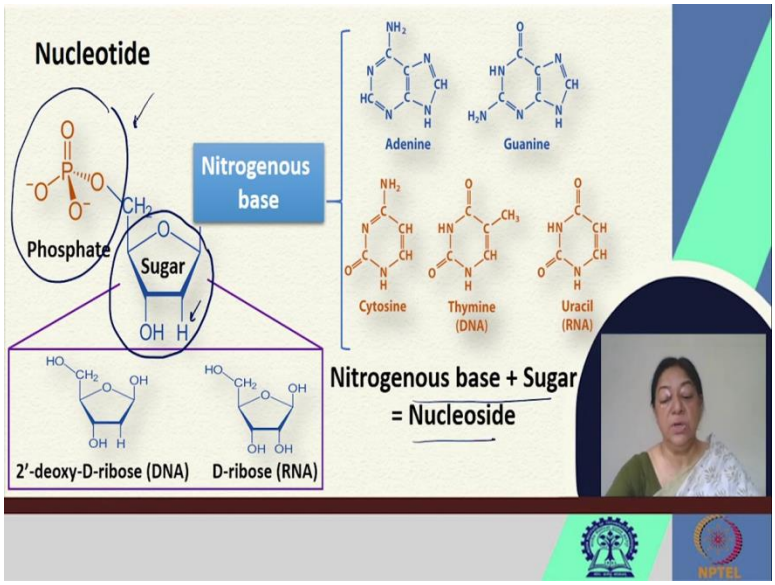
(Refer Slide Time: 04:25)



The two sugars that are involved are the 2'-deoxy D-ribose and the D-ribose present in RNA.

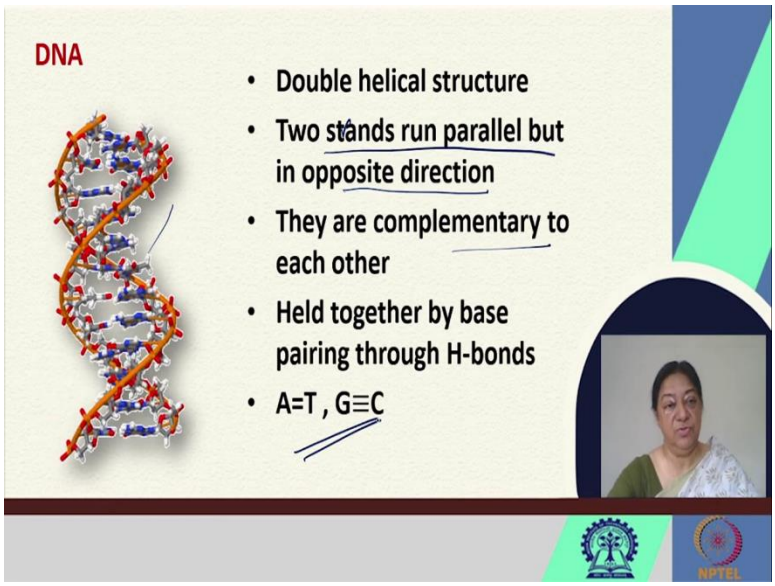
So, the recurring units of DNA contain the pentose sugar and those of RNA contain the ribose sugar. This [refer to slide] is where we have the difference in the sugar of the two nucleic acids; DNA has the 2'-deoxy D-ribose and this is D-ribose for the RNA.

(Refer Slide Time: 05:18)



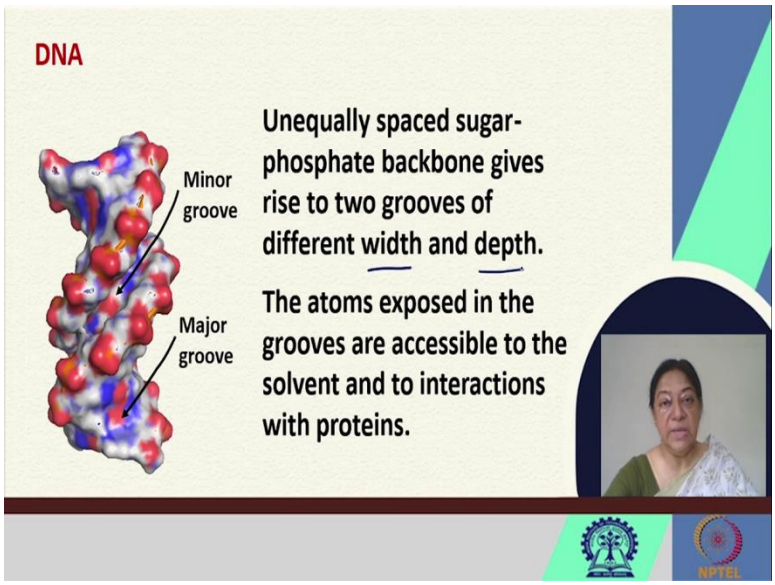
In the formation of the nucleotide, we have the sugar, we see [refer to slide] that the OH is missing here. So this is a deoxy sugar,with the oxygen is missing. We have the nitrogenous base attached here and here is the phosphate group. These three components give us what is called a nucleotide. The nitrogenous bases as we saw could be the purines or the pyrimidines and the sugars could be the deoxy ribose or the D-ribose. The nitrogenous base and the sugar alone are called the nucleoside. When the phosphate is attached, it is called the nucleotide.

(Refer Slide Time: 06:03)



The structure of DNA is unique in its double helical array, where the sugar phosphate forms the backbone and the nitrogenous bases form specific hydrogen bonds between them. The two strands actually run parallel in opposite directions, they are complementary to each other and they are held together through adenine thymine, 2 hydrogen bonds and guanine cytosine 3 hydrogen bonds.

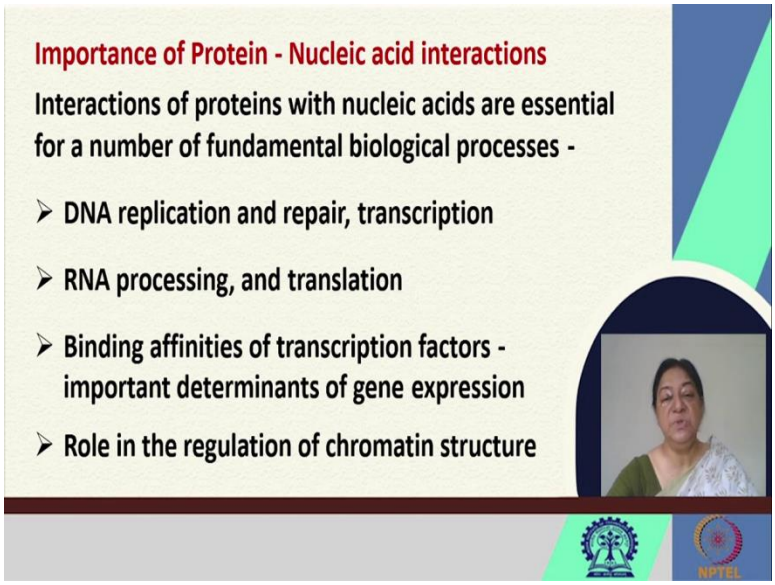
(Refer Slide Time: 06:41)



We look at the structure of DNA, we find that the unequally spaced sugar phosphate backbone gives rise to two grooves that have different width and depth. These are called the major groove and the minor groove and because of the complementarity in the structure of smaller molecules or even with proteins, there will be different molecules that would be capable to bind to these different grooves of the DNA structure.

The atoms that are present here are accessible not only to the solvent, but also to interactions with other molecules, including proteins.

(Refer Slide Time: 07:30)

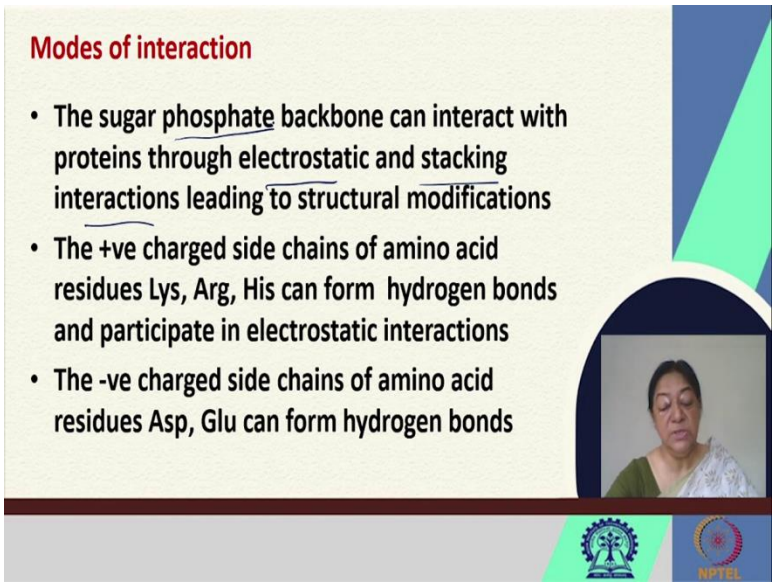


With an understanding of what nucleic acids are in a very brief background, we understood that we had the sugar phosphate backbone and to it was attached the nitrogenous bases. Now, if we look at the importance of the protein nucleic acid interactions, they are extremely important for

several fundamental biological processes. One is DNA replication and repair in transcription, RNA processing and translation, the binding affinities of several different transcription factors that are important determinants of gene expression and the role in the regulation of the chromatin structure.

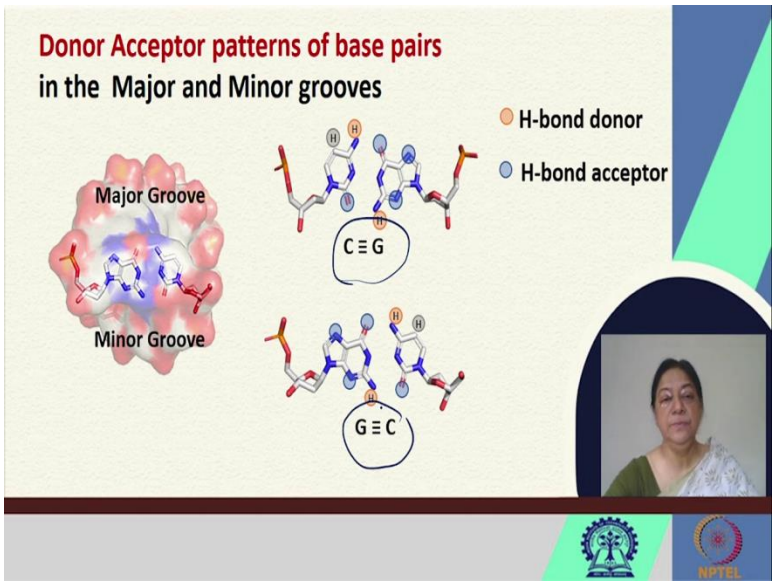
So the protein nucleic acid interactions as we can see, are extremely important to life processes and the types of interactions possible are important. If we look at the modes of interactions, the sugar phosphate backbone can interact with the proteins through electrostatic and stacking interactions because of the phosphate group here.

(Refer Slide Time: 08:31)



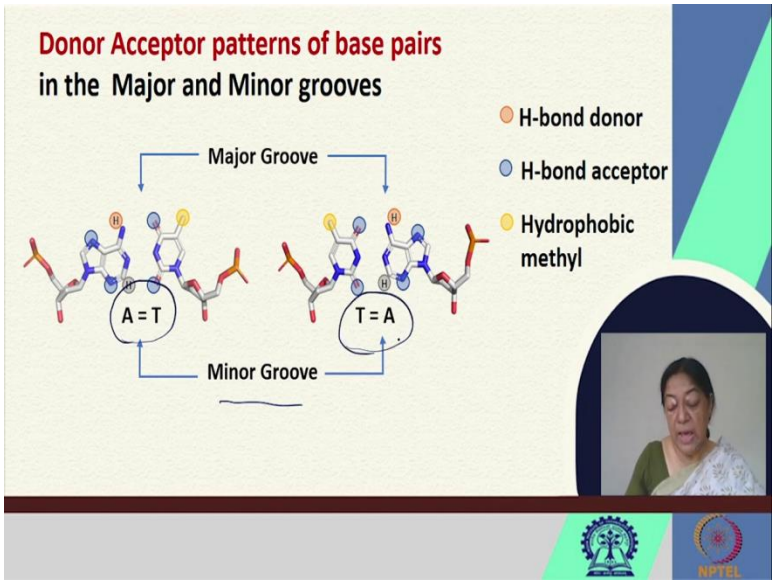
The positive charge side chains of amino acid residues present on the protein, can form hydrogen bonds and participate in these electrostatic interactions and the negatively charged side chains of the amino acids, aspartic acid and glutamic acid can be involved in a hydrogen bonds as well.

(Refer Slide Time: 09:14)



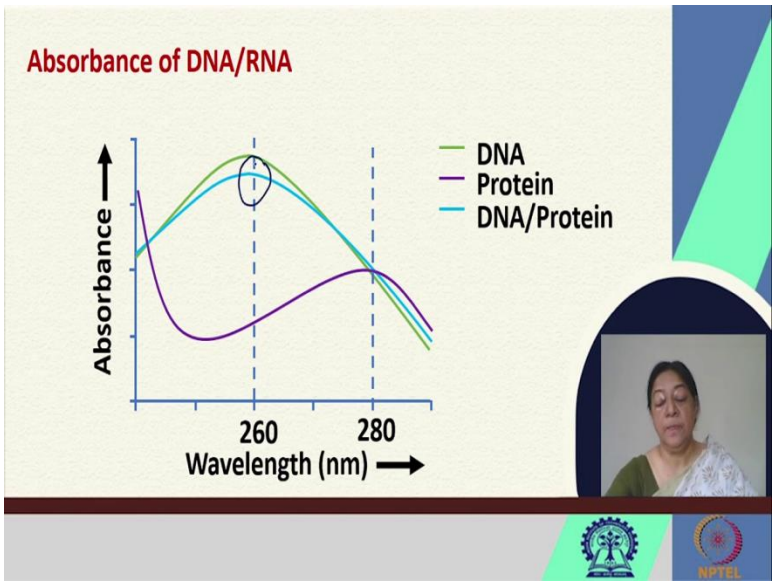
If we look [refer to slide] at the donor acceptor patterns of the base pairs in the major and the minor grooves in this case, we have the specific orientation of the bases where we can have hydrogen bond donors and hydrogen bond acceptors in the formation of the 3 hydrogen bonds that are seen between these two specific nitrogenous bases; the cytosine and the guanine.

(Refer Slide Time: 09:45)



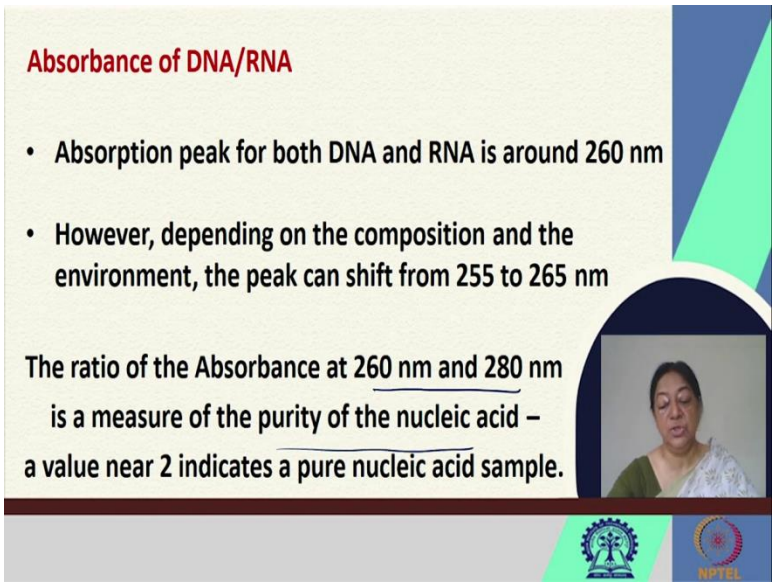
This is important in the stability of the DNA molecule as well. When we look at the donor acceptor base patterns for the minor group set, we will have an AT double bond formation and a hydrophobic methyl that is also present.

(Refer Slide Time: 10:02)



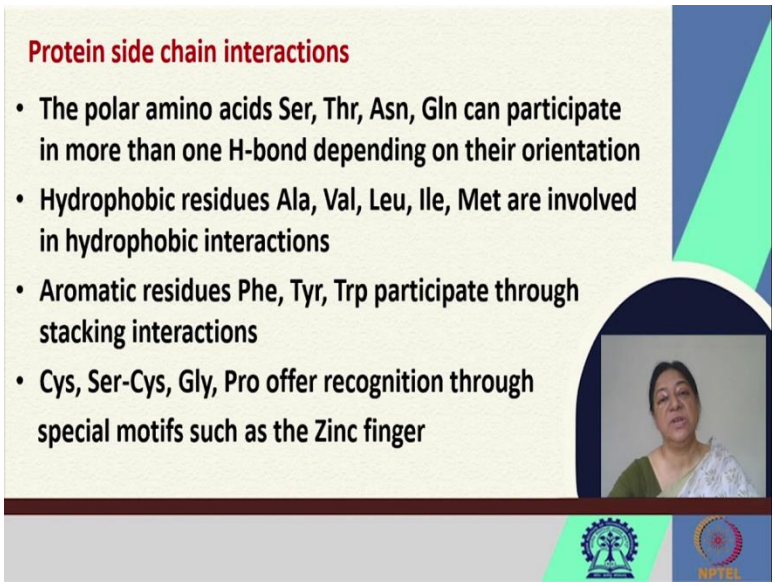
The identification of DNA and RNA is done by absorbance studies. As we had visited absorbance values and absorbance spectra of proteins, we can look at their factor with DNA. This [refer to slide] is where the bases of the DNA and the RNA absorb at a value of 260 nm and due to the presence of the aromatic amino acid residues in proteins, this absorbance is at 280 nm. However, if we have a DNA protein complex, then we have an absorbance around 260 nm here.

(Refer Slide Time: 10:51)



The absorbance peak for both DNA and RNA is therefore around 260 nm. However, depending on the composition and the environment the peak can shift and the ratio of the absorbance at 260 nm and 280 nm, is the measure of a purity of the nucleic acid sample and a value near two indicates a pure nucleic acid sample. This is often a way to identify the presence of the nucleic acids in methodologies that are used for the preparation or isolation of nucleic acids.

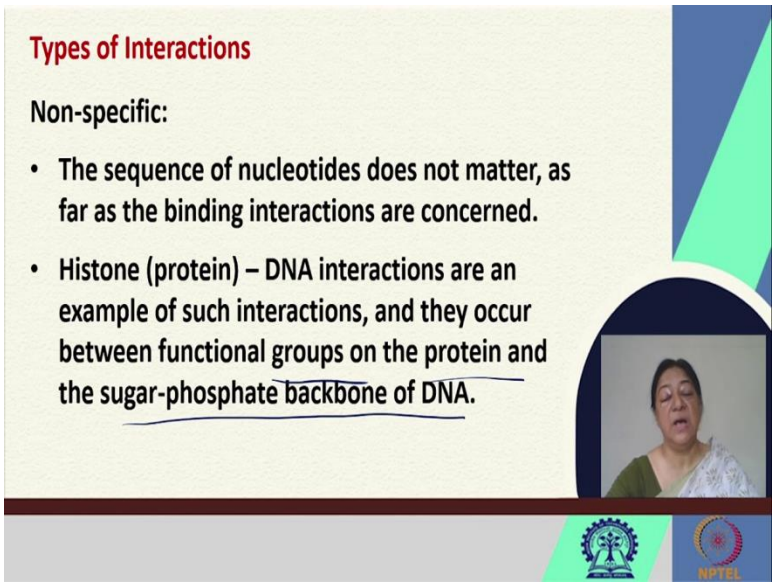
(Refer Slide Time: 11:29)



If we look at the protein side chain interactions that are involved, we have the polar amino acid ser, thr, asn and gln, that can participate in hydrogen bonding depending upon their orientation in the protein, the hydrophobic residues ala, val, leu, ile and met, that can be involved in hydrophobic interactions.

The aromatic residues phe, tyr and trp, that are important here can participate through stacking interactions and cys, ser-cys, gly and pro, offer recognition through special motifs. We will visit the protein and the DNA in their interactions, in the subsequent lectures.

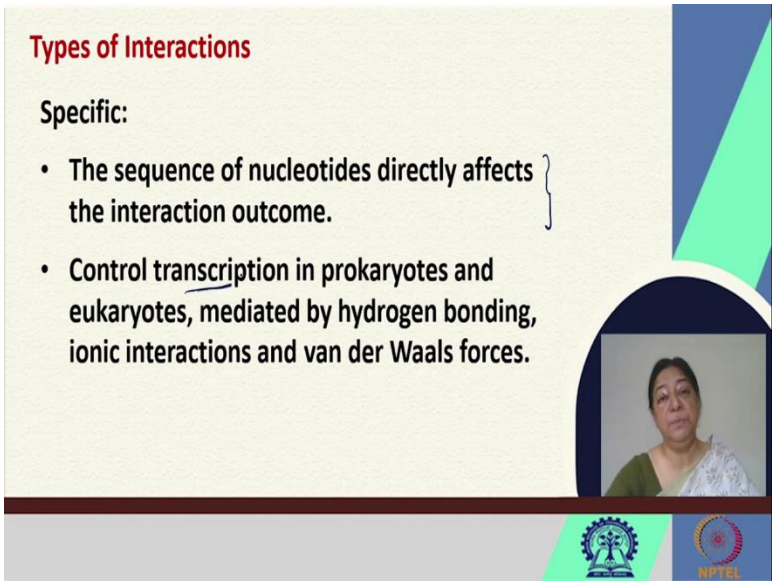
(Refer Slide Time: 12:08)



The types of interactions that we see, can be non-specific in nature. What we mean by non-specific is, the sequence of the nucleotides does not matter as far as the binding interactions are concerned. So, binding will occur irrespective of what sequence the nucleotides are in. The histone protein, DNA interactions are an example of such interactions and they occur between the functional groups on the protein and the sugar phosphate backbone of DNA.

We know that the sugar phosphate backbone of DNA is negative. So, we would like to look at positive functional groups on the protein. These are the specific amino acids.

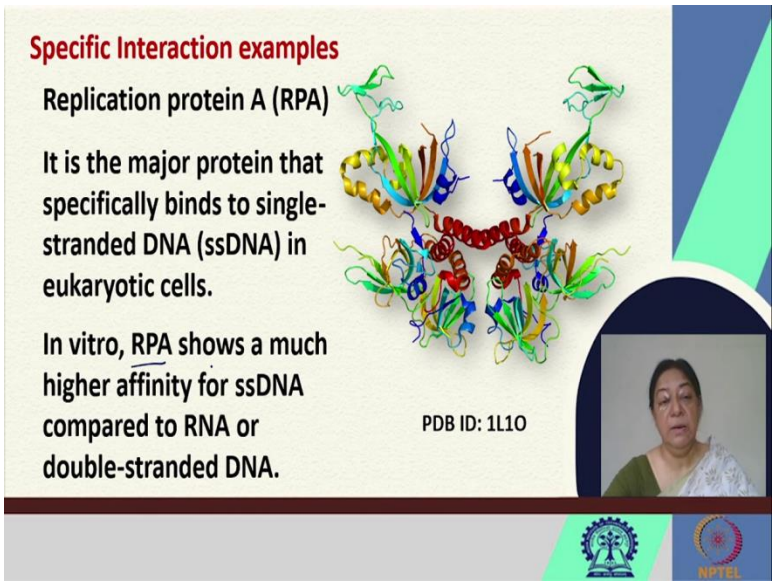
(Refer Slide Time: 12:59)



However, if we look at specific types of interactions between the nucleic acids and protein the sequence of the nucleotides directly affects the interaction outcome and this is important to control transcription in prokaryotes and eukaryotes that are mediated by hydrogen bonding, ionic interactions and van der Waals forces.

So, the interactions between the specific amino acids and specific sequence of nucleotides, are important for the control of transcription and other specific processes that are involved.

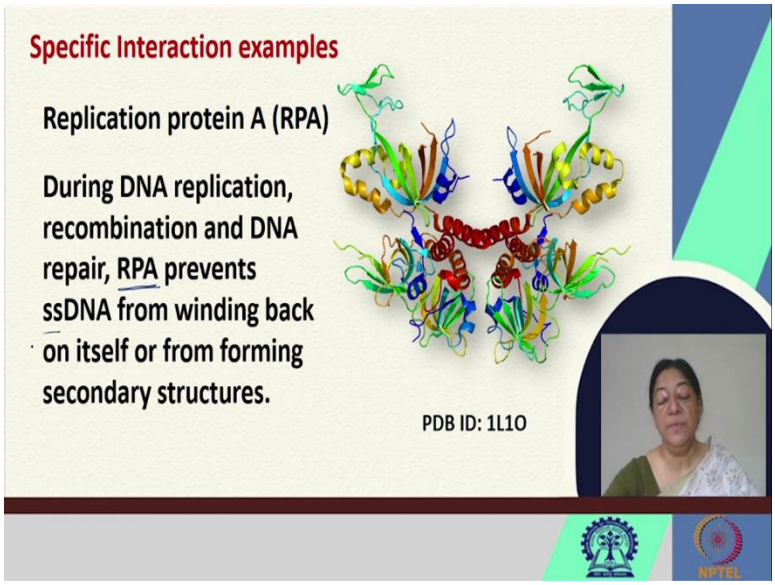
(Refer Slide Time: 13:47)



The specific interaction examples are the replication protein A (RPA). This is a major protein that specifically binds to single stranded DNA in eukaryotic cells. In vitro what happens is, RPA that is the replication protein A shows a much higher affinity for the single stranded DNA, compared to RNA or double-stranded DNA.

So the association here, is important for the replication process and there is a specific sequence that is recognized by the protein, the replication protein A, that binds to the specific sequence of a single stranded DNA in the binding process.

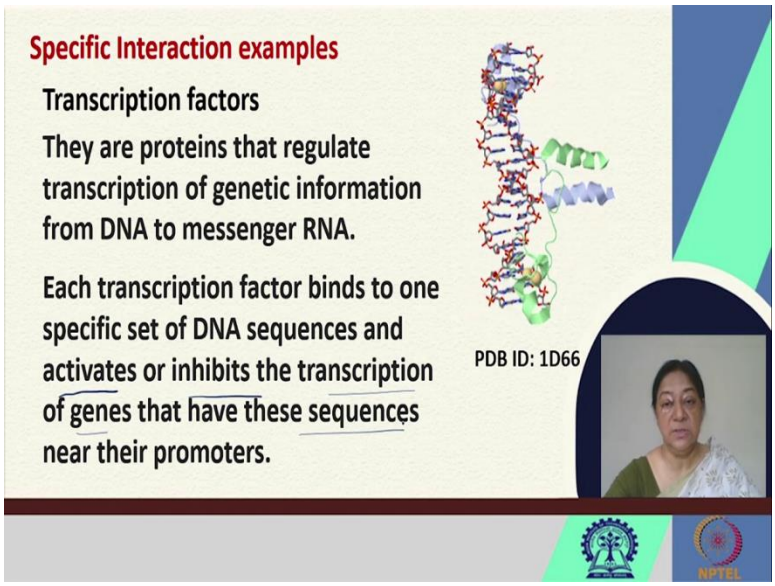
(Refer Slide Time: 14:40)



During DNA replication, recombination and DNA repair, what happens is the RPA will prevent the single stranded DNA from winding back on itself in forming other types of secondary structures. What this does is, it helps in the DNA replication process by binding specifically to

single stranded DNA. In the event, it prevents the single stranded DNA from winding back on itself and forming other types of secondary structures.

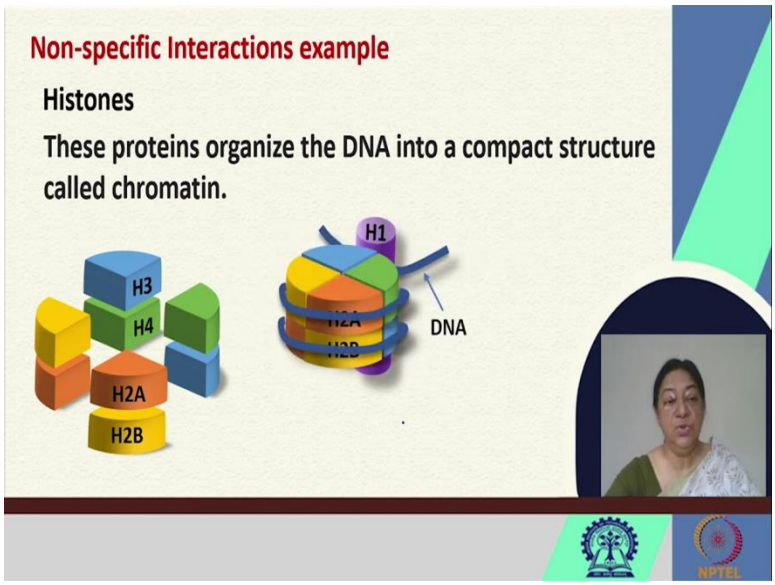
(Refer Slide Time: 15:20)



Other specific interaction examples include transcription factors and these are proteins that are going to regulate the transcription of genetic information from DNA to messenger RNA and this itself indicates the importance of the process because it is going to be the transcription process and specific interaction is required to bring about this. So, each transcription factor will bind to one specific set of DNA sequences only and in the process, it will activate or inhibit the transcription of genes that have these sequences near the promoters.

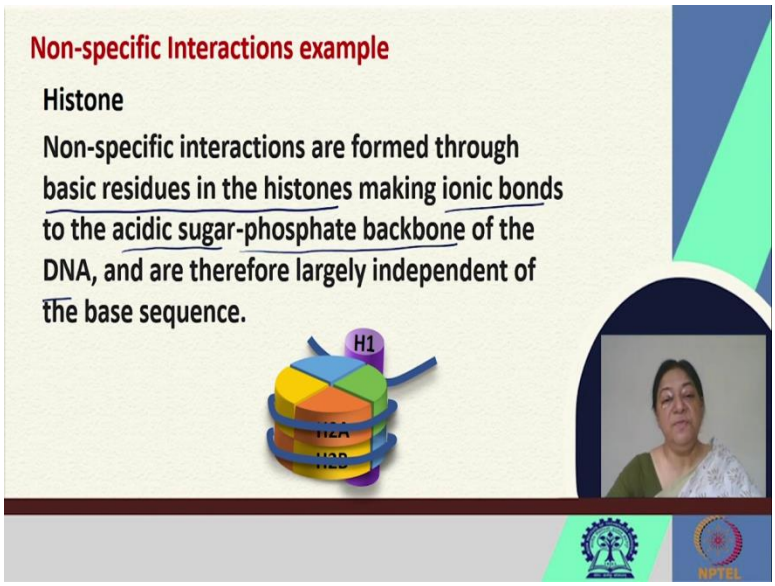
So, the transcription factors will bind to a specific set of DNA sequences only, to bring about the biological process in its functionality.

(Refer Slide Time: 16:23)



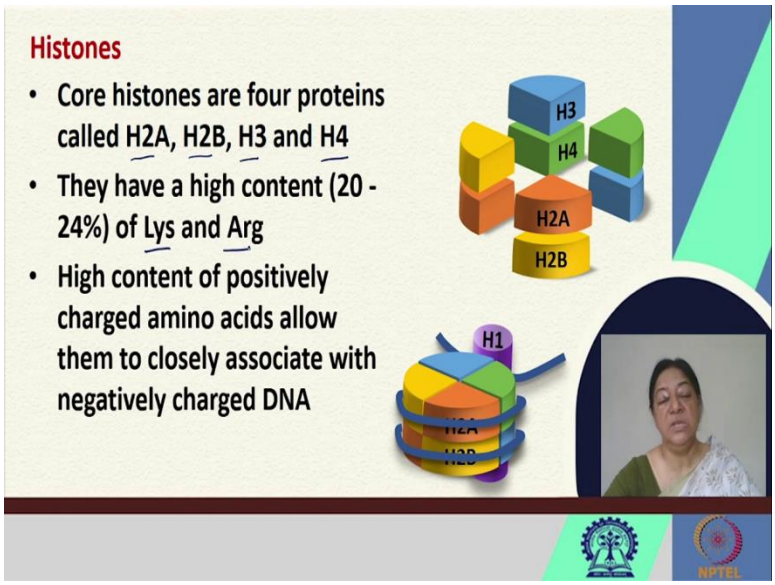
The non-specific interaction examples include the histones. These proteins organize the DNA into a compact structure, that is called the chromatin.

(Refer Slide Time: 16:42)



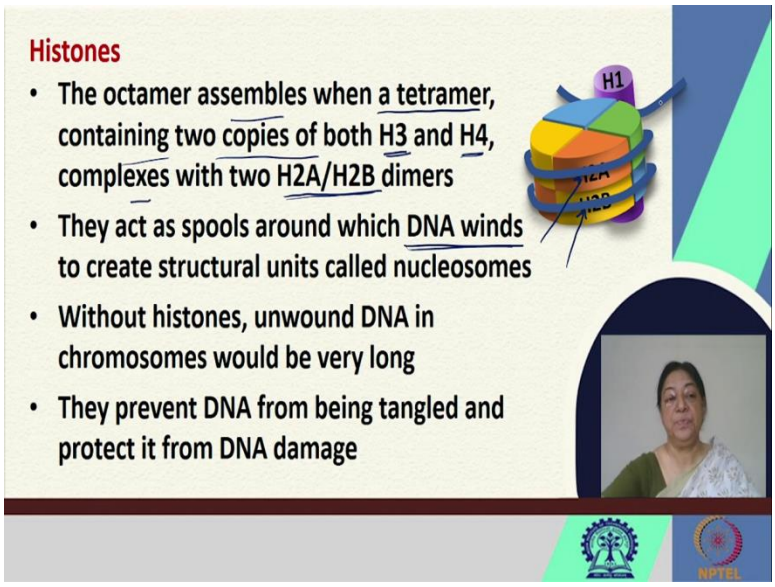
These are the histones that have specific structural aspects; the histone itself is an example of a non-specific interaction and these are formed through the basic residues in the histones that make ionic bonds to the acidic sugar phosphate backbone of the DNA. They are largely independent of the base sequence because they bind preferably to the sugar phosphate backbone. The sugar phosphate backbone being negatively charged, their protein interactions with specific amino acids are important.

(Refer Slide Time: 17:15)



The core histones are four proteins that are referred to as H2A, H2B, H3 and H4 histones. They have a high content of lysine and arginine as would be expected because they interact specifically with the sugar phosphate backbone of the DNA, which we know is negatively charged. This high content of the positively charged amino acid residues, allows them to closely associate with the negatively charged DNA in non-specific interaction types.

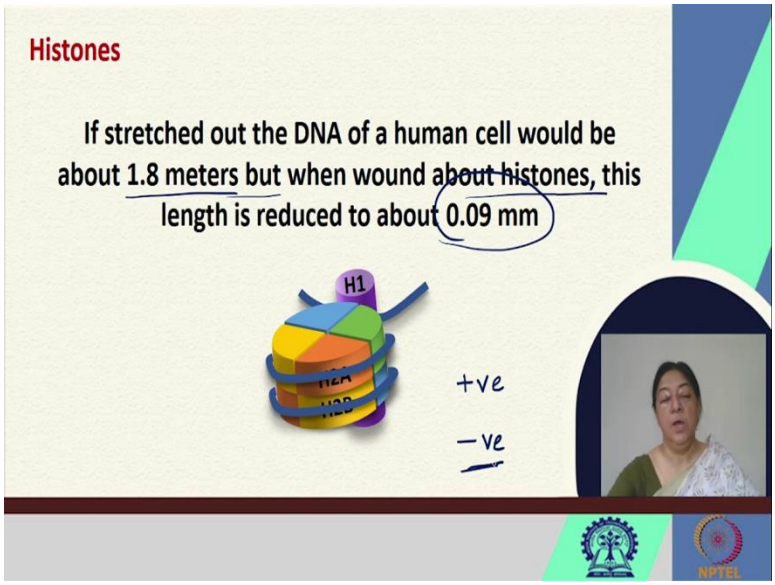
(Refer Slide Time: 18:08)



The octamer of the histone, assembles when a tetramer containing two copies of both the H3 and the H4, forms complexes with two H2A and H2B dimers. So, we have an octamer assembly. This protein is rich in positively charged amino acid residues and after the assembly of the tetramer, they act as spools around which the DNA winds as we can see [refer to slide] the DNA strands that are marked here, winding behind the octamer that is formed.

Now, without the histones we realize that the unwound DNA in chromosomes would be very long. So, the important role of the histones is to prevent DNA from being tangled and protect it from DNA damage.

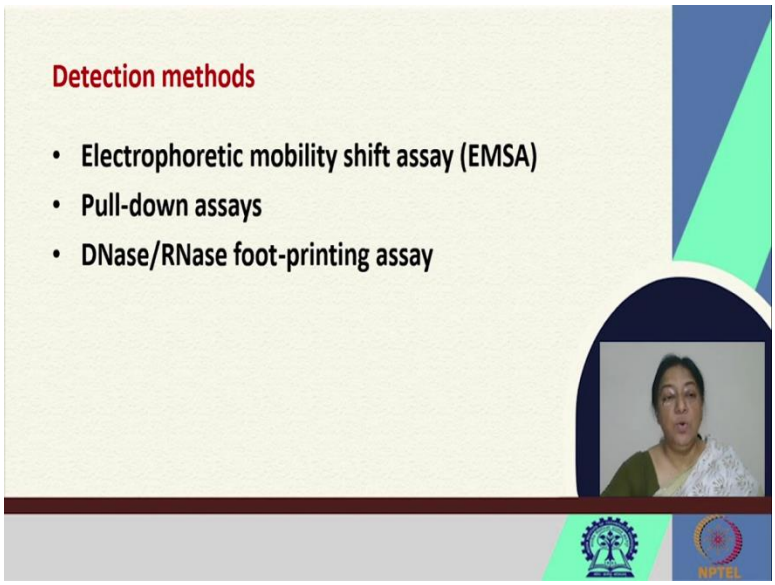
(Refer Slide Time: 19:11)



If stretched out the DNA of a human cell would be about 1.8 m long, but when wound about the histones, this length is reduced to about 0.09 mm. So, we realize the importance of these specific interactions such as these histones, that are largely populated with lysine and arginine residues up to 20% to 24 %; that give us an idea about how the positively charged residues on this histone protein are going to interact with the negatively charged sugar phosphate backbone of the DNA.

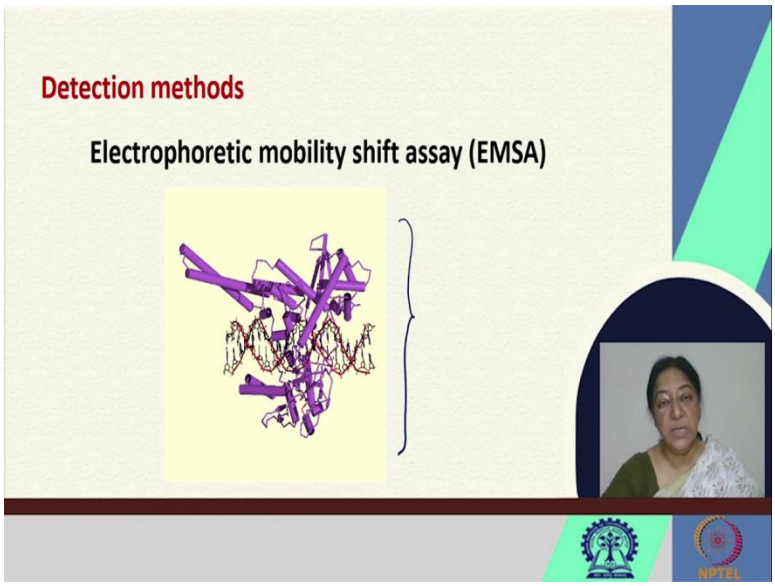
So we realize the importance of the protein and nucleic acid interactions. They may be specific in nature particularly when there is a transcription or a translation process involved, that takes us from DNA to RNA to protein. These interactions or the biological processes that involve these proteins, can be specific in nature, can be non-specific in nature, depending on the special mechanism that they are involved in.

(Refer Slide Time: 20:36)



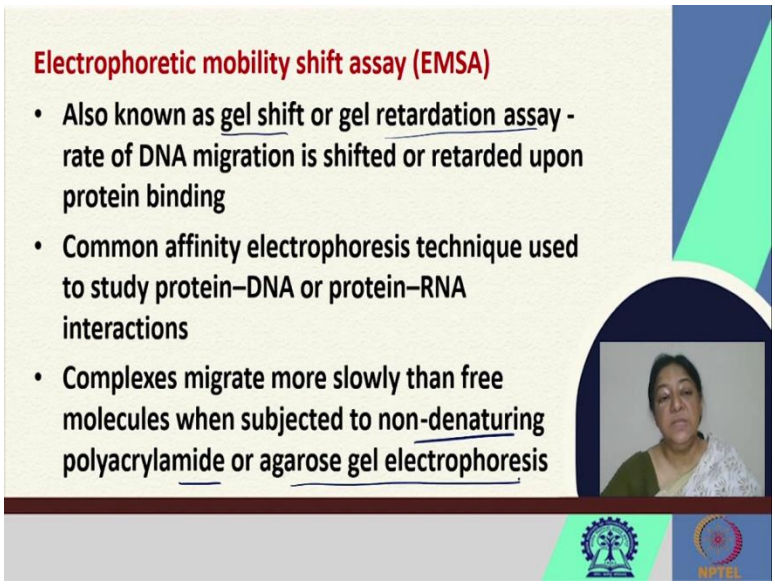
We now look at specific detection methods associated with protein nucleic acid complex formation. In the common detection methods that are used, there is the electrophoretic mobility shift assay, the EMSA as it is called, the pull down assays and the DNase or RNase foot-printing assay. In each of the three lectures that we have in this module part for protein nucleic acid interactions, we will be looking at one of the detection methods.

(Refer Slide Time: 21:15)



In this lecture we will be looking at the electrophoretic mobility shift assay, with respect to DNA protein complex formation. Given that we have our DNA protein complex form, we would now like to assess its presence and to see whether this protein nucleic acid complex has formed. This may be specific in nature or non-specific in nature.

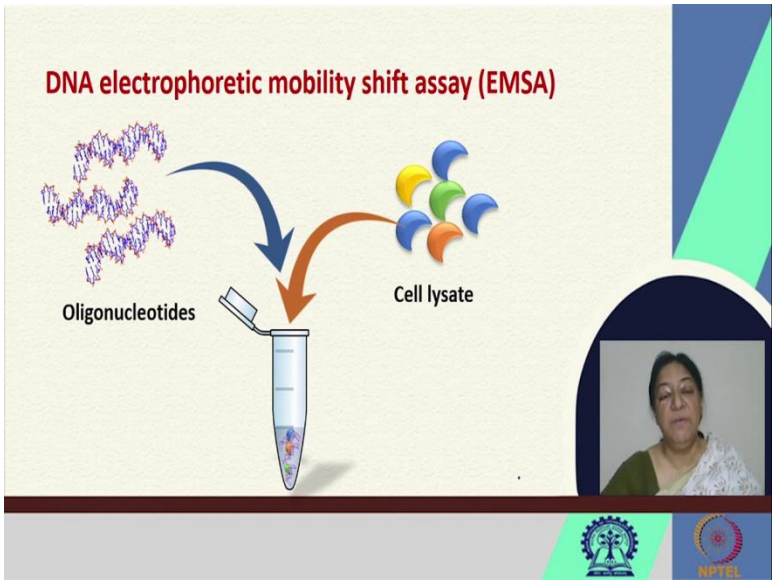
(Refer Slide Time: 21:41)



In this specific assay, we have the gel shift or the gel retardation assay, as it is also called. We are monitoring the rate of DNA migration that is shifted or retarded upon protein binding. What essentially happens is, a common affinity electrophoresis technique is used to study these protein DNA or protein RNA interactions. On complex formation, the free molecules when subjected to the non-denaturing or the polyacrylamide agarose or gel electrophoresis, we realize that on complex formation the migration is going to be much slower than the free molecules.

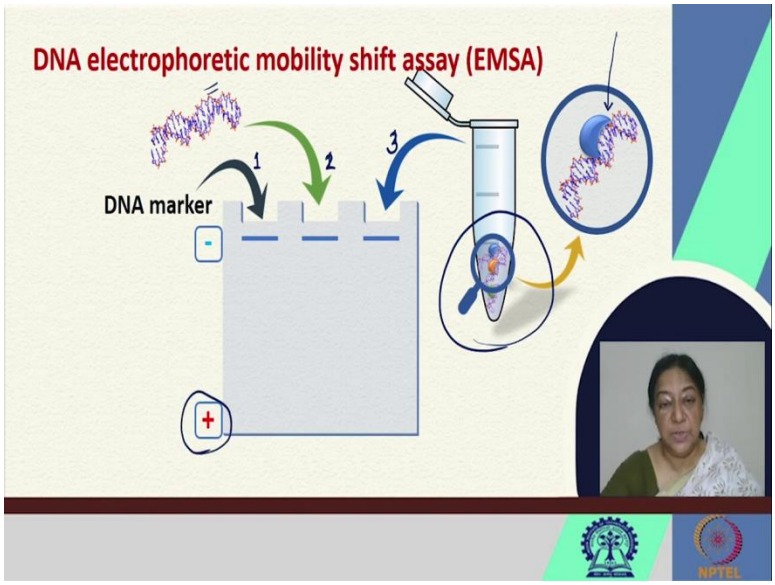
Considering a regular gel electrophoresis experiment, a non-denaturing type in this case, whether it is a polyacrylamide or agarose gel electrophoresis, we will be able to monitor the migration of the bands of the free molecule and the complex.

(Refer Slide Time: 22:54)



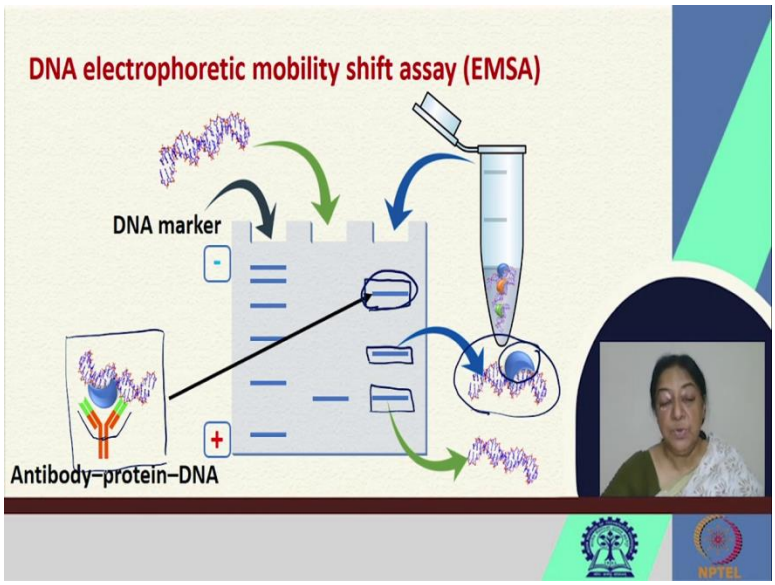
So we have our oligonucleotides and we have our cell lysate that contains our protein of interest. These are mixed together in our eppendorf tube and given the affinity for DNA, for the specific protein, we will have a DNA protein complex formation.

(Refer Slide Time: 23:18)



This [refer to slide] DNA protein complex formation as is shown here, is present in our eppendorf tube along with free proteins and free DNA. The first lane that we see here lane number 1, is a DNA marker that gives us an idea of the specific molecular weights that we may observe. The second lane corresponds to DNA alone and the third is our mixture. So we have number 2 here and number 3 here. This is our complex mixture that contains our DNA protein complex of interest. When this is subjected to electrophoresis, to an electric field; we will have migration.

(Refer Slide Time: 24:10)

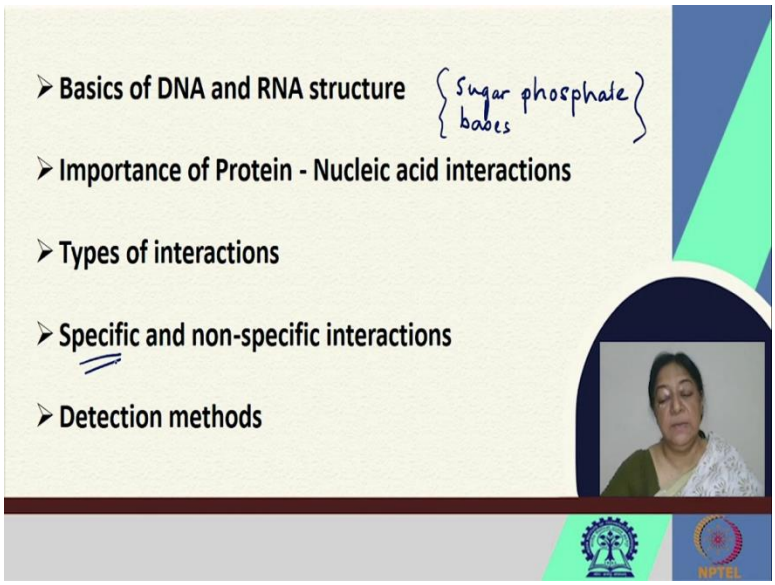


This migration is going to occur based on the size of the complex formed. In this [refer to slide] case we will see that the retardation or the slower migration is because of the complex formation of the protein with DNA. The second band or the faster moving band that we observe here is due to DNA alone that corresponds with our DNA present here. So this is our DNA protein complex, this is our DNA.

Another way in which this [refer to slide] may be even retarded further, is by the formation of an antibody protein DNA complex, where the specific antibody that we have is for the protein of interest. So, we have an even larger complex now, due to the presence of the large antibody and what is going to happen is, the migration is going to be even slower and our band will appear even further away from this.

So, the migration that we see would depend upon the size of the molecules. Initially the first band observed here is the uncomplexed DNA, the second band we observe here is the complexed DNA with the protein of interest and the third band that we observe here is due to the antibody protein DNA complex. This is often called the super shift assay because we have now an antibody protein DNA complex which results in even further retardation, further slower migration of our protein DNA complex.

(Refer Slide Time: 26:09)

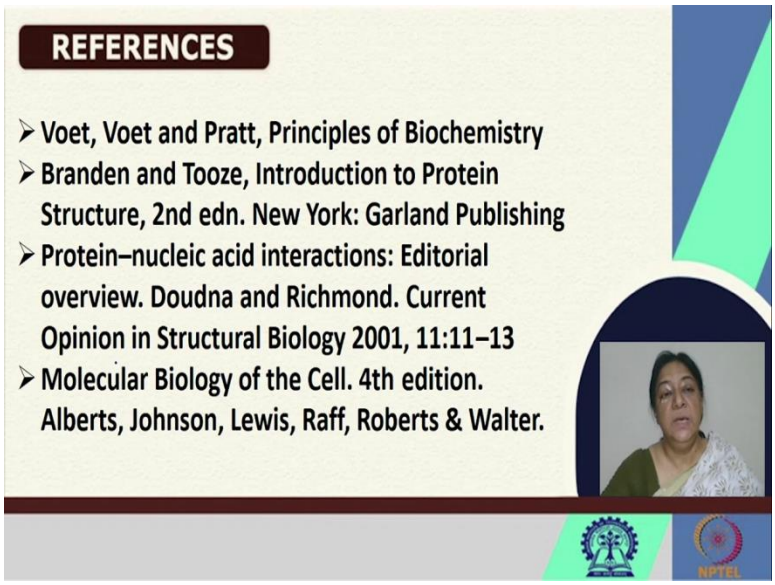


What we have looked at in this lecture, is the basics of DNA and RNA structures following from the sugar phosphate backbone. Then we looked at the specific purine and pyrimidine bases and how they form the DNA and RNA structure. Our interest here is in the importance of protein nucleic acid interactions, considering our central dogma of biology from DNA to RNA to protein and understanding protein nucleic acid interactions is extremely important, in understanding many processes.

The types of interactions depend upon the specific amino acid residues involved. These may be hydrophobic type, small polar type that would be involved in hydrogen bonding or electrostatic interactions that are most common. These interactions can be specific in nature, where the sequence of the DNA is important in the protein nucleic acid interaction process or they could be non-specific interactions like we looked at the specific example of the positively charged histone proteins, that interact with the negatively charged DNA.

In the detection methods for this lecture, we looked at the electrophoretic mobility shift assay that gives us an idea of the formation of the complex of the protein nucleic acid in an identification method, to determine whether the complex formation has taken place.

(Refer Slide Time: 27:55)



These [refer to slide] are the references.

Thank you.