Experimental Biochemistry
Prof. Soumya De
School of Bioscience
Indian Institute of Technology, Kharagpur

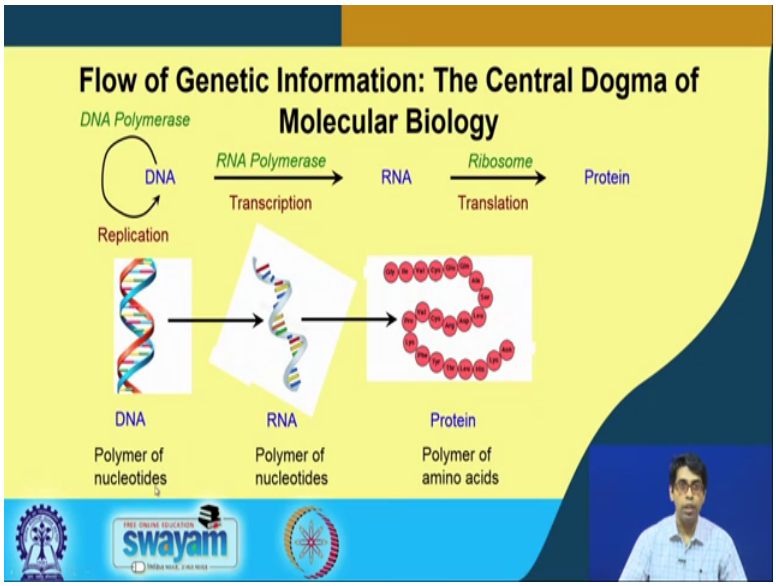
Lecture - 38
Isolation and Characterization of DNA Part – I

Welcome to week 9. So, in today's lecture we are going to talk about Isolation and Characterization of DNA. In a previous lecture that is in week 7 we talked about isolation and characterization of proteins and today we are going to look at isolation and characterization of another biological macromolecule which is DNA.

(Refer Slide Time: 00:41)

So, in today's class we are going to talk about DNA which is the genetic material and how to separate that from other cellular components just like protein. And, we will see that this isolation depends on the type of DNA and also from the source and size of the DNA.

(Refer Slide Time: 00:59)



So, before we go into the actual experimental details let me step back and talk about the basic central dogma of molecular biology. So, here what we see is the central dogma of molecular biology which was first proposed by Sir Francis Crick who along with James Watson proposed the double helical structure of DNA. So, what is the central dogma of molecular biology? It tells us the flow of information that DNA is the genetic material which stores all the information that is necessary to carry out all the functions in the living system that is the cell or in organism.

So, DNA quotes the information and on the other extreme we have protein, which carry out all the function. So, proteins like in the previous lecture we talked about enzymes, enzymes are catalysts then there also other types of proteins which are not enzymes, but then they also carry out different functions like receptors on cell surfaces or there are cytoskeleton proteins. So, there are so many different types of proteins which carry out most of the functions inside a cell. So, this is the functional part or the functional macromolecule and this is the macromolecule which stores the information.

In lecture 7 we saw how we can isolate proteins and do characterization. In characterization using various bio physical methods; today we are going to look at DNA. So, DNA stores the information when a cell divides the DNA also has to multiply. So, that copying of DNA is called Replication and it is done by enzymes which are again proteins these enzymes are called DNA Polymerase.

Now, once you have the DNA if you want to make the protein you have to extract out the information. So, when we talk about DNA we are talking about the full DNA for example, in case of bacteria let us say E-coli the DNA is several million base pair long. So, typically it is around 5 million base pair long, but that long DNA has information for different types of proteins.

So, if you want to make one particular protein that would that information might be encoded in only 1000 base pair long stretch of DNA also that particular part which is referred to as a gene is called copied into a temporary copy called the RNA. So, both DNA and RNA are nucleotides. So, will see in more details what we mean by nucleotides. So, the small part of DNA is copied into RNA and this process is called Transcription and it is carried out by another type of enzyme called the RNA polymerase.

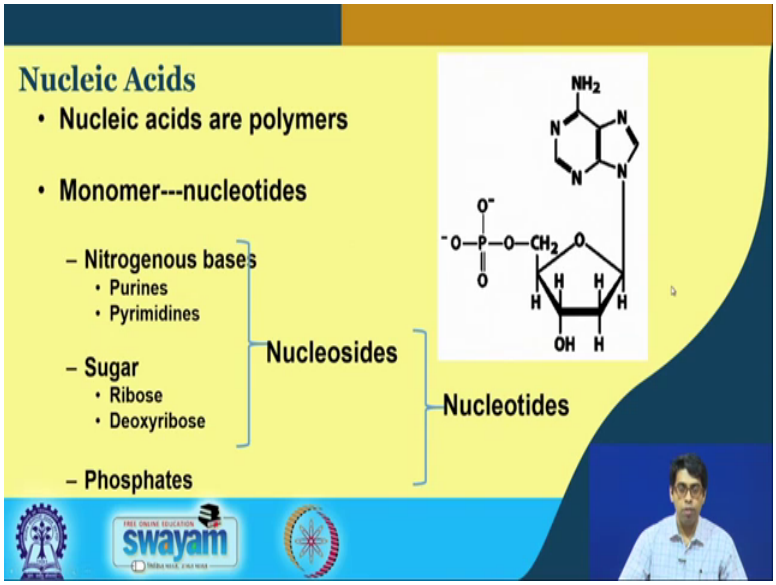
So, this small part of nucleotide this small nucleotide has the same information as that of the DNA, but this is only a small part. So, this are this is only 1000 base pair long which encodes information for only one protein then that information is read by another machinery called the Ribosome to synthesis protein and this process is called Translation. We have seen that once the protein is formed it is a linear chain of amino acid. So, the protein is a polymer of amino acids and then this polymer spontaneously folds, similarly DNA and RNA are also linear polymers and they are polymers of nucleotides. So, this is the different type of monomer different from the amino acids.

Again there are 20 different amino acids in case of proteins, but there are only 4 different types of nucleotides in case of DNA and RNA. One major difference between DNA RNA and Protein is in the shape of this molecule. So, we have seen that once the linear protein molecule is synthesized it is spontaneously folds into a 3 dimensional structure. So, we look at that in the lecture where we studied protein folding and unfolding. DNA on the other hand has a very unique structure. So, it does not depend too much on the sequence of the DNA, all the DNA molecules fold up into a structure which is called the Double Helical Structure.

So, typically most DNA have this double helix structure. If you see carefully you will see that there are 2 polymer chains which are intertwined with each other. So, that is why it is a double helix RNA is a single chain. So, that is a difference between DNA and RNA and will see some more differences and we are going to exploit those differences to

purify DNA from RNA and proteins, because in a cell all of these macromolecules are present at any given point of time.

(Refer Slide Time: 06:07)



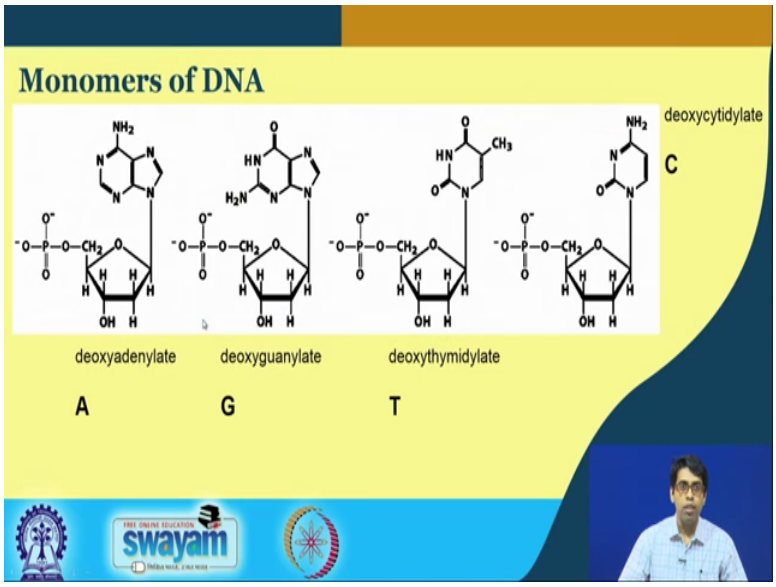
So, what are Nucleic acids? This is the typical structure of a nucleic acid. So, these nucleic acids are polymers and this is the monomer of a nucleic acid and these monomers are called Nucleotides. If you see there are 3 different components in this nucleotide, the first component is the nitrogenous base. So, this is the nitrogenous base and there are 2 types of nitrogenous base, Purines and Pyrimidines and again will see in the next slide there are 2 purines and 2 pyrimidines. So, there are 4 different types of nitrogenous base which is shown up here.

Then the second component is the sugar moiety. This ring like structure is the sugar moiety depending on whether this is a DNA or a RNA the sugar is deoxyribose or ribose. The difference between these two sugar is in this carbon. So, this sugar molecule has 5 carbons this is 1, this is 2, this is 3, this is 4 and this is 5; at the second position you will see that both are hydrogen atoms. So, this is deoxyribose. In case of ribose this H which is at the bottom this H is OH. So, that is the ribose sugar and when this oxygen is removed it becomes deoxyribose

The third component is the phosphate group which is attached to the fifth carbon of the ribose or deoxyribose sugar. These two when you have only nitrogenous base and the sugar. So, the nitrogenous base and the sugar if you have only this 2 and not the

phosphate then it is referred to as a nucleoside. When you also have the phosphate present at the 5 frame position then this monomer is referred to as a Nucleotide. So, in today's lecture we are only going to refer to nucleotides.

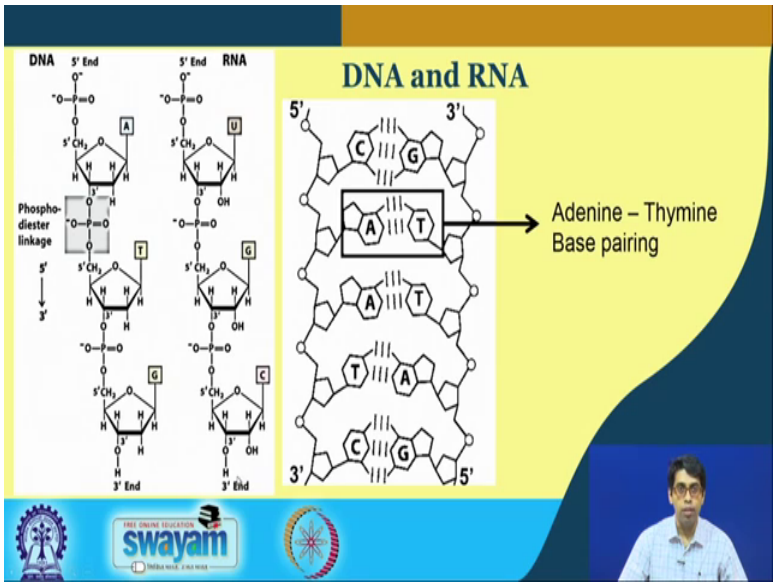
(Refer Slide Time: 08:21)



So, now coming to all the 4 different nucleotides that is the monomers of DNA. So, the one that we saw in the previous slide was this which is known as Deoxyadenylate or in short A. So, we just like the proteins were we have 20 different monomers and each of the monomers where given 20 different letters from the English alphabet; similarly in case of DNA the 4 monomers are giving 4 single letter codes the first one is Deoxyadenylate and it is referred to by A, the second one is Deoxyguanylate you see that the base is the one that is changing and it is referred to by the letter G.

Third one is Deoxythymidylate or T and the fourth one is Deoxycytidylate or C. If you see carefully the sugar and the phosphate groups are exactly the same in all this 4 nucleotides it is only the base which is attached at the first position of the sugar this base is different. So, it is only the nitrogenous base that is different and between all this 4 monomers.

(Refer Slide Time: 09:39)



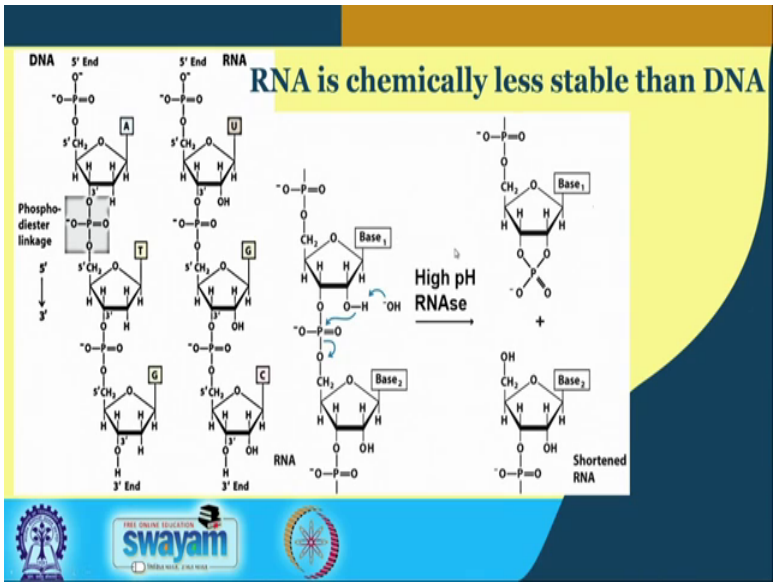
Now, these monomers are attached just like the proteins where the amino acid monomers are attached to form a polymer in case of DNA and RNA these nucleotide monomers are attached to form a linear polymer. So, on the left hand side you see this linear polymer is that of DNA and on the right hand side this linear polymer is that of RNA. So, if you compare them you will see that the difference between these two is in the nitrogenous base. So, here the second position has H and here the second position has OH in case of RNA. So, this is the ribosugar and this is the deoxyribosugar, another difference that you will see is in the bases will come to that later. So, how does this polymer form.

So, this is the monomer the base nitrogenous base attached to the first position the ribose deoxyribose sugar and then the phosphate group attached to the fifth position. Now you see that another phosphate group is attached to the third position. So, the third OH and this is connected to the fifth CH 2 OH of the next nucleotide. So, this forms the phosphodiester linkage between 2 monomers. Similarly there is another phosphodiester linkage between these 2 monomers and it proceeds so on and so forth. This phosphate is attached to the 5 frame carbon. So, this is referred to as the 5 frame end this phosphate or this OH. So, this is the 3 prime carbon and this is referred to as the 3 prime end. So, just like a protein where we have the N terminus and C terminus in case of DNA or RNA we have the 5 prime end or the 3 prime end.

So, I was referring to another point that apart from these 2 prime position the nitrogenous bases are also different. So, A, G and C are exactly same in case of both DNA and RNA, but only the T is something that is different. So, instead of thymidine we have uracil in case of RNA. So, it is one of the bases that is different and this 2 prime position where it has a OH in case of RNA and H in case of DNA. Another difference between DNA and RNA is that DNA forms a double standards structure. So, it is shown here this is one of the strands. So, you see it runs from 5 prime to 3 prime direction and this is another strand which runs in the opposite direction.

So, this is 3 prime and 5 prime and what happens is the base C and base G they can form hydrogen bond similarly base A and base T they can form hydrogen bond. So, these hydrogen bonds, these pattern holds this 2 strands together and we have a double stranded DNA and we will see in a the next few slide that this strands forms a spiral like structure which is referred to as the double helical structure. This type of double helical structure is absent in RNA because RNA mostly they are single stranded.

(Refer Slide Time: 13:33)

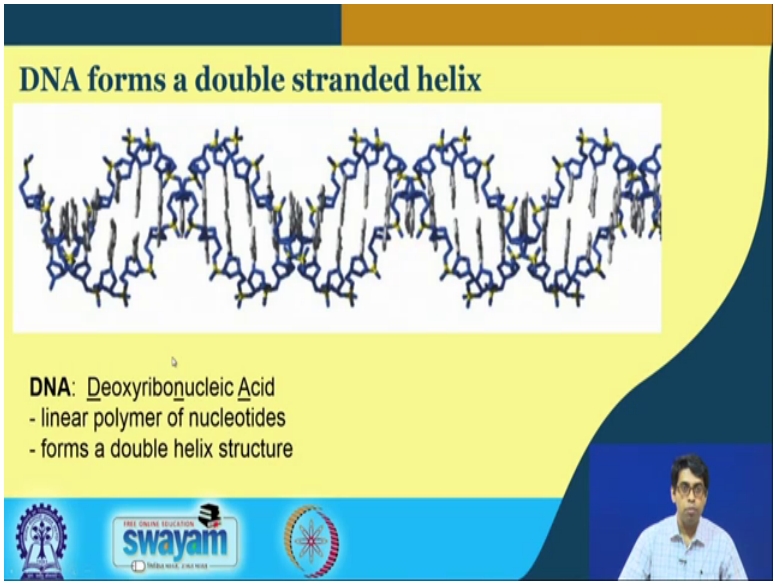


This is something that is very important specially for today's lecture that RNA it has this OH at the 2 prime end and the presence of this OH makes it chemically less stable compared to DNA because the presence of this OH results in a particular reaction where a hydroxyl ion which can be supplied by high ph or it can be also enzyme catalyzed this

hydroxyzline can attack this so that there is a nucleophilic attack and this phosphodiester linkage is broken. So, that now you have the breakage in the RNA strand.

So, this type of reaction cannot happen here because we do not have an OH group here. So, this makes DNA molecules much more stable chemically compared to RNA more of a double stranded nature of DNA makes it even more stable compared to RNA. So, we are going to exploit this particular difference in the chemical nature of DNA and RNA when we are going to separate the DNA and RNA molecules.

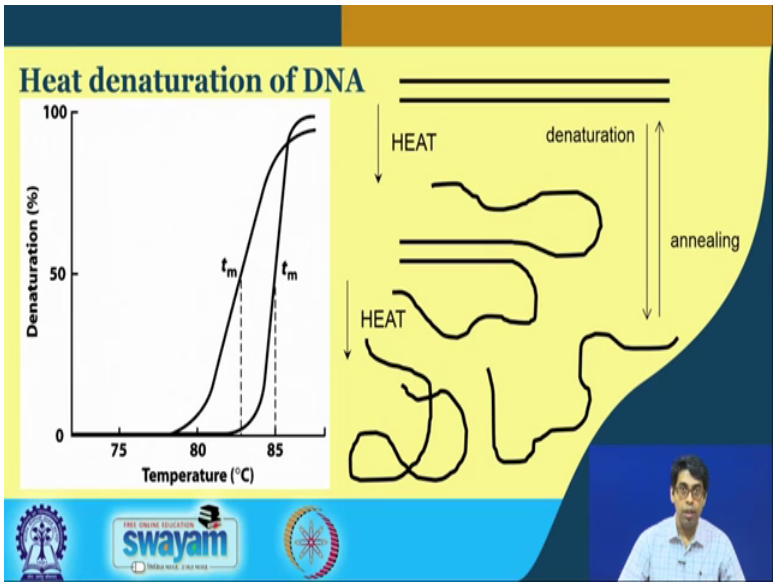
(Refer Slide Time: 14:51)



So, as I mentioned before the 2 strands of DNA can form a double helical structure like this. So, you can see that one strand is this you see it is helical structure. So, it is spirals and the other one follows it and forms another spiral. So, together they form a double helix and the nitrogenous bases are all inside whereas, the sugar and phosphate which forms the backbone are on the outside.

So, this results in a highly negative charge on the DNA molecule because all the phosphate groups carry one negative charge. So, deoxyribonucleic acid or DNA is a linear polymer of nucleotides and 2 such polymers come together to form a double helical structure.

(Refer Slide Time: 15:47)



Another important properties of DNA which again we are going to exploit during it is purification is the heat denaturation this is very similar to what we have seen in case of protein. So, in case of protein if you heat them the protein unfolds and if you cool it then the protein can again go back reversely to form the folded protein we see something very similar in case of DNA. So, the 2 strands of DNA are represented here these 2 lines are the 2 strands of DNA which are held together by hydrogen bonds between the nitrogenous bases which are on the inside.

If I heat a solution of DNA what will happen is these hydrogen bonds will start breaking. So, the 2 strands will sort of peel away from each other. So, the more heat I give the more hydrogen bonds will be broken. So, that at some point of time that all the hydrogen bonds will be broken so that the 2 strands of DNA will be completely separated from each other. Now if I cool this solution again this is the reversible process. So, it can go back and form this double stranded DNA. So, the breaking of hydrogen bond and separation of the 2 strands is called denaturation and formation of the hydrogen bond and formation of the double helical structure upon cooling down the solution is called annealing of the DNA molecule.

So, it can be a reversible process and this denaturation and annealing also depends on the length of the DNA and composition of the DNA because in the previous slide slides we saw that GC pairs from 3 hydrogen bonds whereas, 80 pairs form 2 hydrogen bonds. So,

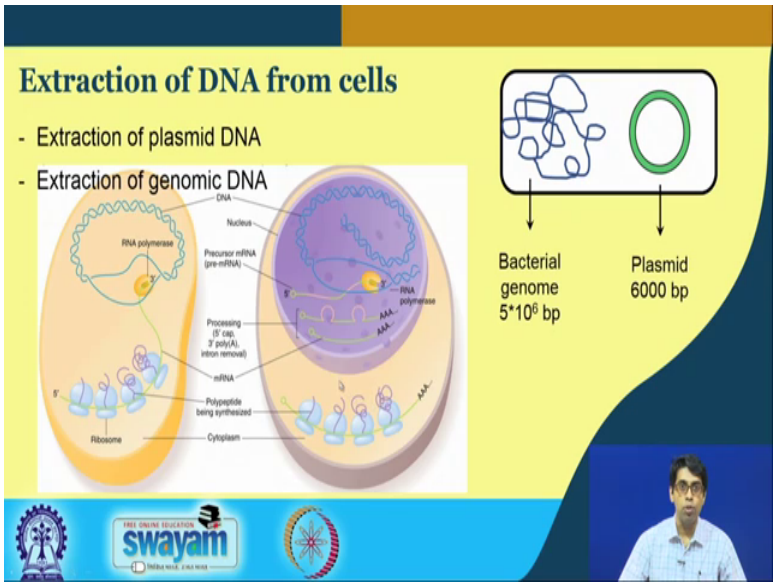
depending on how much GC content is there in the DNA the amount of heat that will be required to denature a particular double stranded DNA will increase and that can be easily measured by a very simple property called the melting temperature.

So, what we are plotting here is. So, suppose we have taken two different DNA solutions let us say the first one is A and the second one is B. When we heat them and we observe the absorbance of this DNA. So, the absorbance will change because the absorbance of it when it is double stranded and when it is completely isolated strands will be different. So, let us say the absorbance increases. So, as the DNA molecule denatures more and more of the molecules are denaturing the absorbance increases and when it is completely denatured all of it will be. So, the will reach a saturation point.

So, for two different DNA molecules or two different sequences we see two different curves the midpoint where we reach fifty percent denaturation. So, this is nothing, but the middle point between the saturation point and this base line. So, if we take as the fifty percent denaturation the temperature at which it reaches this point is referred to as the melting point and we see that the melting point is different for the 2 different DNA molecules. So, this melting point will depend on the length of the DNA molecule and it will also depend on the sequence of the 2 DNA molecules.

So, we are going to exploit this fact that DNA can denature and it can also be annealed back. What we are not going to do is we are not going to heat our solution we are going to use chemicals to do the denaturation just like we saw urea and polydium hydrochloride can be used to denature proteins we will see that the same chemicals or a different chemical can also be used to denature DNA molecules.

(Refer Slide Time: 20:01)



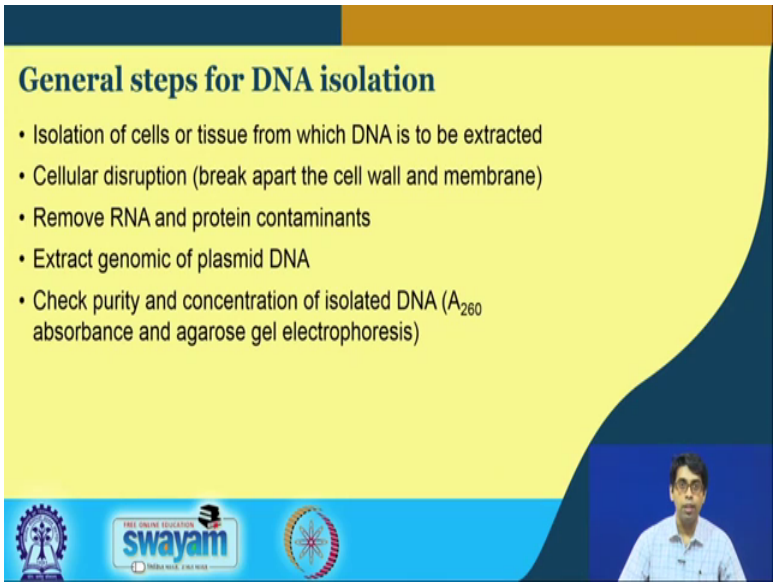
So, in this week we are going to see extraction of DNA from cells and will mostly talk about extraction of DNA from bacteria. So, in case of bacteria there are two different types of DNA one is called the plasmid DNA and the other one is the genomic DNA. So, genomic DNA is the DNA which encodes all the proteins and other information for the normal function of the bacteria or the survival of the bacteria. So, it is it is own DNA and typically this DNA if we talk about aquiline will be something in the order of 5 times ten to the power six base pair law.

So, that is 5 million base pair law. Plasmid DNA are small circular DNA and they are much smaller. So, the typical size of a plasmid DNA will be in the range of six thousand base pair law. So, I am not going to talk much about plasmid DNA in this week's lecture because we are going to cover in lot in more details about we are going to talk in more details about plasmid DNA in next week's lecture when we talk about recombinant DNA technology. So, this is something that we can artificially introduce into a bacteria in order to make certain proteins.

So, sometimes we might require to purify only the plasmid DNA and some other times we want to we may want to purify the genomic DNA and another small information that, so this is a bacterial cell. So, it is a prokaryotic cell and this is a eukaryotic cell what we want to see here is that if this is the cell you see that in case of bacteria the DNA is here and all the proteins are here. So, everything is in the same place. So, when we break open

this cell all of them will be out in the solution and from there we have to selectively purify the DNA in case of eukaryotes the DNA is in the nucleus, but again when we break open the cell membrane we break open everything and all the cellular components are out in the solution and from there we have to purify the DNA.

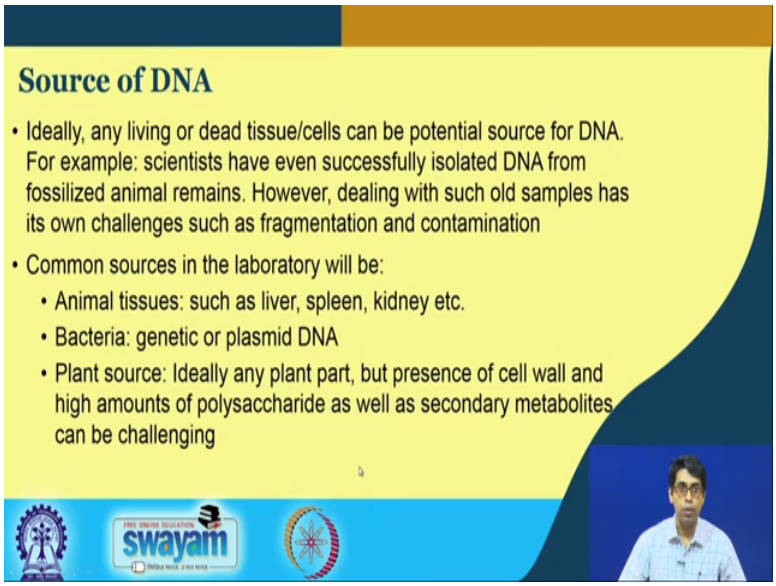
(Refer Slide Time: 22:33)



So, it is quite a challenging task and we have to go through several steps to get a very high quality of pure DNA. So, what are the general steps the first step as we saw in case of protein purification is that you have to isolate the cells. So, in case of protein purification we were working with bacterial cells in this case also we can use bacterial cells, but if it is some mammalian cells then we will have to isolate the cells from the tissue from which the DNA has to be extracted. Once we have this cells we will disrupt the cell membrane in case of bacteria we also have to break up what the cell wall. So, that everything that was inside comes out into the solution.

Now, this solution will have DNA, RNA, protein and all sorts of cellular derby. So, we have to selectively remove RNA, we have to selectively remove all proteins and we have to also remove all the cell membrane components and other molecules and finally, we will extract the DNA the genomic or the plasmid DNA. Once we have our DNA we have to make sure that it is pure. So, we will do gel electrophoresis to check the purity of the DNA and we will also check the absorbance at 260 nanometers to check the concentration which will tell us the amount of DNA that we have when able to purify.

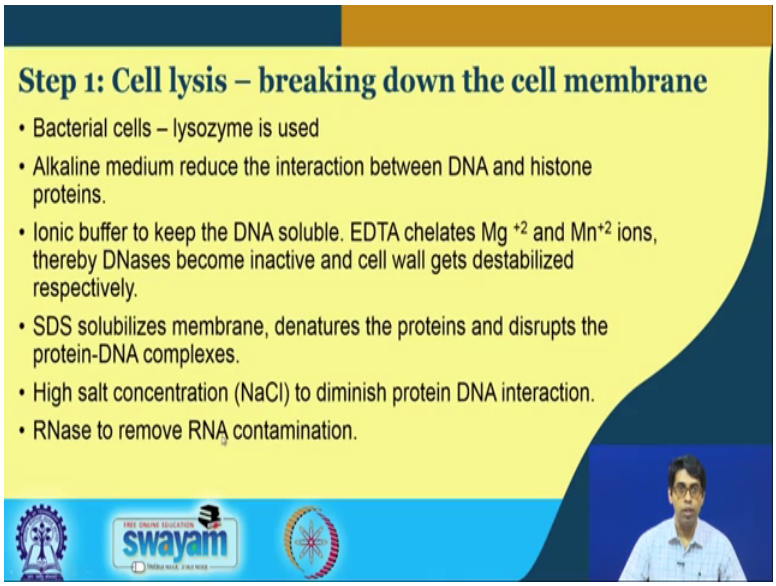
(Refer Slide Time: 24:03)



So, what is the source the source of the DNA can be anything. So, it can be any living or dead tissue and normally we have to extract out the cells from the tissue. So, these are all potential source for the DNA and the. So, here is one example where scientist have been able to successfully isolate DNA from fossilized animal remains because as I mentioned that DNA is a very stable molecule, but depending on the quality of the DNA it can be isolated. So, in case of laboratory the common sources are different animal tissues such as liver, spleen, kidney etcetera and in case of bacteria we take the whole. So, we can grow the bacterial cells in the culture and depending on our requirement we can purify the genetic or the plasmid DNA from the bacteria.

In case of plants we can also isolate DNA, but it poses some extra challenges. So, any part of a plant can be used, but the presence of cell wall which is a very rigid structure which helps to maintain the shape of the cell in case of plants and this high amount of polysaccharides poses extra problem. So, there are extra steps that you will have to go through to extract the DNA from plant sources.

(Refer Slide Time: 25:45)



So, what are the steps in; so, in this lecture I will just briefly go through the steps and then I will talk about them in more details in the next lecture. So, if we take the example of bacteria, the first cell the first step will be to break down the cell membrane this is something very similar that we did during protein purification. In case of protein purification we sonicated the proteins the bacterial cells and we also used lysozyme. So, in this case we are going to use a very small amount of bacterial cells we are not going to sonicate we are only going to use lysozyme.

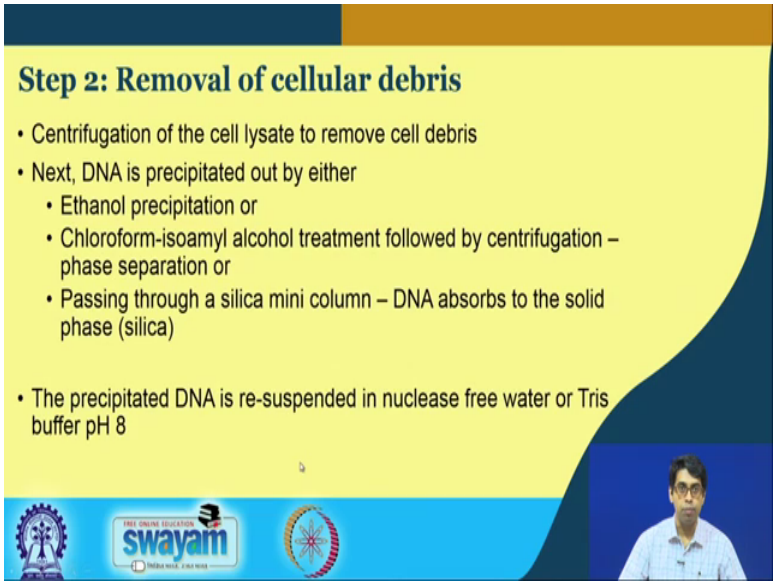
So, lysozyme is an enzyme which breaks down the bacterial cell wall and this is the enzyme which is actually found in our eyes also and it actually protects our eyes from bacterial infection. So, we use the same lysozyme to break open the bacterial cells. Once everything is out in the solution then what we need to do is we have to selectively separate DNA from proteins because, the DNA can be associated with different types of proteins. There are transcription factors, there are different proteins which bind DNA and most of this interactions are charged interaction, also we have to denature the DNA. So, one of this is achieved by reducing or increasing the pH of the solution. So, what we do is we make a we use a alkaline medium to break the all the hydrogen in hydrogen bond interactions.

We also increase the salt content so, that these different ionic interactions are disrupted and we also use EDTA which is a small molecule that binds different ions such as

magnesium ions, manganese ions. So, these are the ions which are removed so that different enzymes such as DNA's which can break a DNA which can chew up our DNA. So, those enzymes are inactivated, also adding EDTA makes the cell wall of the bacteria unstable. So, it helps in both degrading the bacterial cell wall and also protecting the DNA from several enzymes.

SDS is added to solubilize the membrane. So, this SDS is sodium dodecyl sulphate and it is the same surfactant that we use when we run SDS page gel. So, we have seen there that it can denature proteins. So, when we add SDS it disrupts the membrane and it also denatures the protein so, that all the protein DNA complexes that are present will be disrupted. High salt is used to again disrupt the protein DNA interaction because, this interactions are also driven by electrostatic forces. And finally, we add RNA's to chew up all the RNA that is present in this whole cell lysate solution. Because, we saw in the previous slide that RNA is less stable and RNA's this enzyme selectively degrades RNA molecules. So, the RNA will be chewed up and that will help us to easily remove the RNA molecule.

(Refer Slide Time: 29:23)

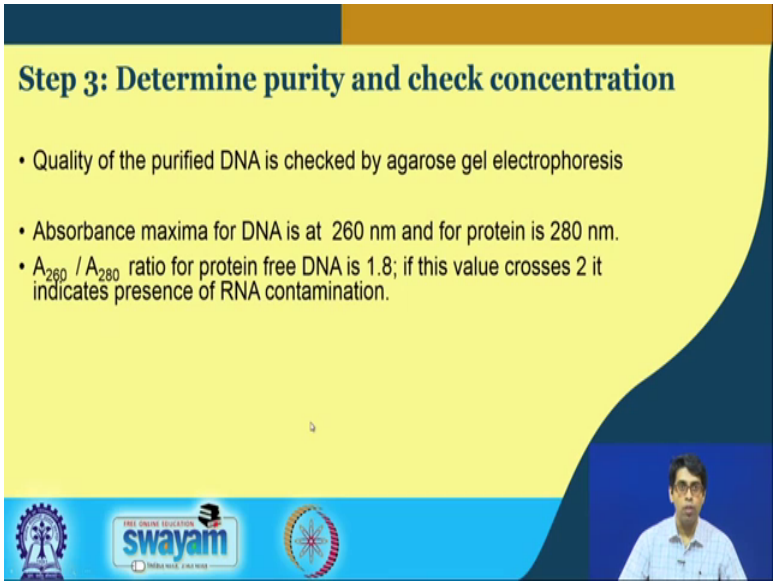


So, once we have all of this, once we have the cell lysate and we have successfully disrupted all this interactions. Now, what happens is that this cell debris the denatured protein, the chewed up RNA they will start to follow aggregate and to speed up the process we put this cell lysate the tube containing the cell lysate into a centrifuge

instrument and we centrifuge. So, that everything is precipitated out and the solution contains the DNA. Now, to purify the DNA and remove it from other small molecules and this monomeric nucleotides, we precipitate the DNA by some non polar solvent. For example, we can use ethanol or we can use chloroform isoamyl solutions to precipitate the DNA and remove the supernatant. and, now we have this white precipitate in these two steps which will be mostly pure DNA. Or, we can take this solution and pass it through a silica gel column which binds selectively binds DNA and all other soluble parts are washed out from the column.

So, either these using any of these two methods we can purify the DNA selectively and get a very high quality of DNA. Once we have the precipitated DNA or we have the DNA bound to the silica column, we can re-suspend it or we can elute it from the column using nucleus free water or some buffer typically a buffer is used which is Tris EDTA buffer pH 8.

(Refer Slide Time: 31:19)



So, now the DNA is in solution we have to just do quality check and determine the amount that we have. So, the quality of the DNA is checked by running it on an agarose gel electrophoresis. Remember that you can run DNA in both agarose and poly polyacrylamide gel, but since these are very big DNA molecules we are not going to use page or polyacrylamide gel electrophoresis; we are going to use agarose gel electrophoresis. And, that if we see single band or very few bands that we can explain we

know that our DNA is highly pure and then you can check the absorbance of this of your sample at 260 nanometer to determine the amount of DNA. What we also do is we also check the absorbance at 280 and then we take the ratio of these two absorbance. If the ratio is close to 1.8 it means that we are getting more absorbance at 260 then 280. So, it means that there is not much protein present in our DNA sample and if the value is less than 2 then it also indicates that there is not much RNA contamination.

So, gel electrophoresis will tell us the purity and absorbance at 280 nanometer tells us the quantity. And also these two ratios, this ratio tells us that there is not much protein and RNA contamination in our DNA sample.

(Refer Slide Time: 32:55)

So, again these are the references and in the next lecture I will go through specific example of extraction of DNA, plasmid DNA from a bacterial source.

Thank you.