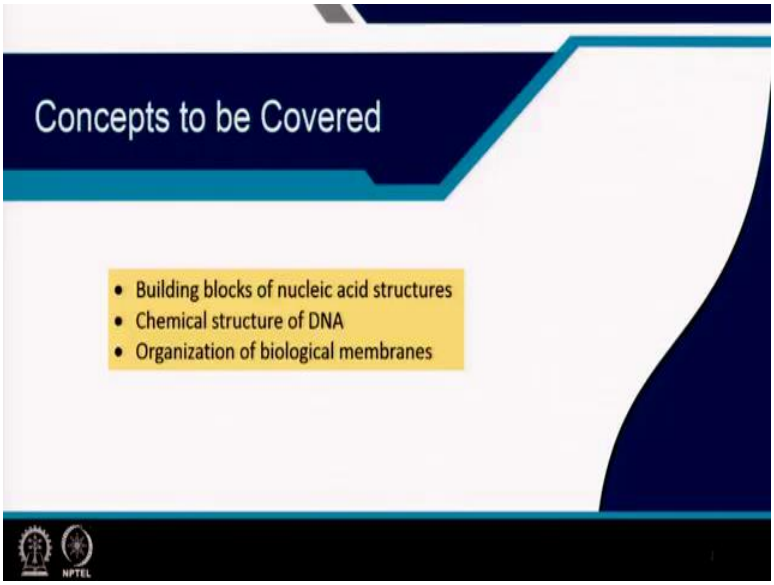
Biological Inorganic Chemistry
Professor Debashis Ray
Department of Chemistry
Indian Institute of Technology, Kharagpur
Lecture - 4
Structures of nucleic acids

Hello welcome back and good morning to everybody. So, the next part of our course on Biological Inorganic Chemistry will be talking about again the module 1, we are still there so lecture 4 is structures of nucleic acids.

(Refer Slide Time: 0:50)



So, one is a this particular one, so why do we see that what are the things we can see now after polypeptide chains, structures and all these now we will see, will bring now another component of this biological world is your nucleic acids. Why we should know about nucleic acids? Because the nucleic acids if we know about the basic component, the molecular structures as well as the final structure like the chemical structures of DNA, we will find there are some useful parts in it, like the nitrogen bases are there, purine and pyrimidine bases we call, not only knowing these structures of this basic purine and pyrimidine, these are the part of the bioorganic chemistry.

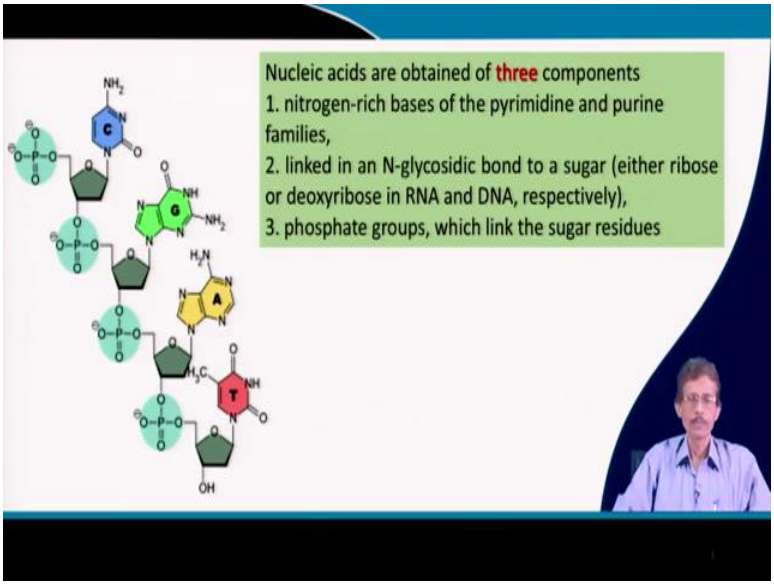
But we should also know that like pyridine we all know the pyridine and the pyridine can form like water molecule can form a coordinate bond to a metal ion center. Similarly, pyridine is not a molecule which we get from the biological world or the nature, from the nature because it is not a part or component of the amino acid residues.

But you can have the other heterocyclic ring like that of your porphyrin which is a pyrrole unit. So, pyrrole nitrogen can be available for coordinate to the metal ion center. Similarly, the histidine residue, the imidazole side chain of that particular histidine amino acid is available we know, we have seen also in our previous classes that the histidine residue can be available for the coordination to the metal ion center.

So, that is why the whole DNA molecule whether the DNA molecules can interact with the metal ion center that is why we should know a little bit about the DNA structures, very briefly, so one particular this half an hour class is devoted for that particular purpose only. And then briefly we will also see the organization of biological membranes because the membranes are there and basically these are some log gates basically.

So, whether that membranes will be permeable to water molecules or they will be permeable to the metal ion centers also, we should know also, whether you can have a huge population of sodium ion outside the cell or inside the cell, what about potassium ion also. So, you think of it and you know in that way and try to understand in that way whether you can have the higher concentration of sodium ion or the potassium ion within the cell or outside the cell.

(Refer Slide Time: 3:25)



So, you can have three components in all these nucleic acids. So, you have the nitrogen rich bases just now I told you that you can have the pyrimidine and purine bases. So, like your pyridine or imidazole part so these pyrimidine and purine nitrogen can be a very good donor to

your metal ion center and you have N-glycosidic bonds also, so you bring, you have to bring the sugar molecules, ribose sugar or the deoxyribose sugar we all know that the corresponding sugar will give you whether you are getting RNA or DNA.

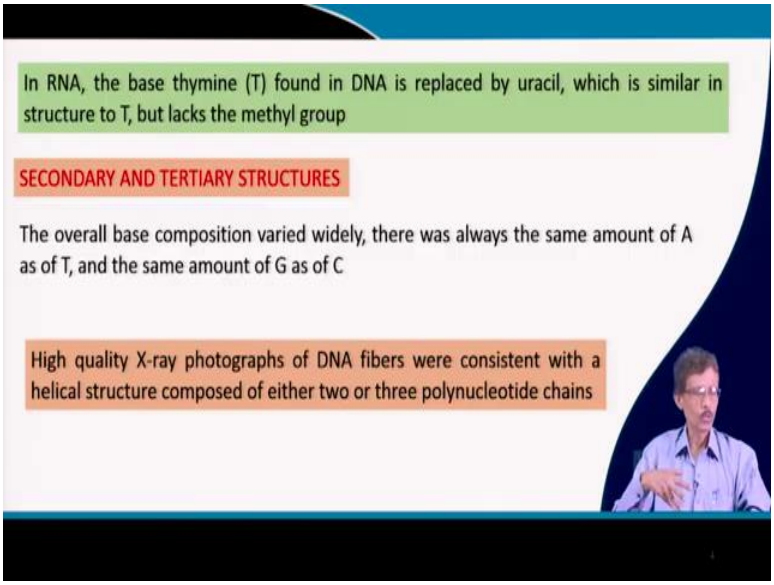
So, ribonucleoside based or the deoxyribonucleoside based you can have for DNA and RNA molecules. Then finally you have the phosphate linkages which link with the sugar residue. So, the typical inorganic part that means that non-metallic part is your phosphate groups.

So, when people can ask you that what biological inorganic chemistry you have learned from this particular course also, related to the corresponding non-metallic part, involvement of the non-metallic species or non-metallic group or the function we all know that the phosphates are typical radicals we call in terms of typical inorganic chemistry nomenclature, in our college days or school days the phosphates you will be able to identify, detect qualitatively as well as quantitatively you can estimate the amount of phosphate present in any other samples, typical in organic samples like phosphate minerals, phosphate ores and all these.

Similarly, from the biological world also you can get many phosphates like this, some of them are inorganic phosphates but some are not when they are attached to the sugar residue they become biological phosphates. So, you see these are the very basic components from these three units.

So, you will have C G A and T are your, the corresponding bases. Then you have the corresponding sugar unit and sugar unit can be of two types, that means the ribose sugar or the deoxiribose sugar, so there is not many complications and then you have the phosphate linkers.

(Refer Slide Time: 5:46)



But how you compare this with DNA and RNA only the base part, so in the base part thymine is found in DNA is replaced by uracil in the structure and lacks the methyl group, so that is the structure how this is related, the T is related to uracil which is U.

So, like your protein structure we can have the secondary and the tertiary structures also, that means much more the higher level of complication can come, primary is the simplest one, secondary little bit more complication you bring then the tertiary structures and all these, the way we say that you have the metal ion, it is binding to a ligand system like porphyrin, you get the metal ion complex but you can have other interactions from the periphery of the ligand part or the bound oxygen molecule in myoglobin through some other non covalent interactions like that of your hydrogen bonding interaction that we consider that your secondary coordinate sphere is getting stabilized always these are important.

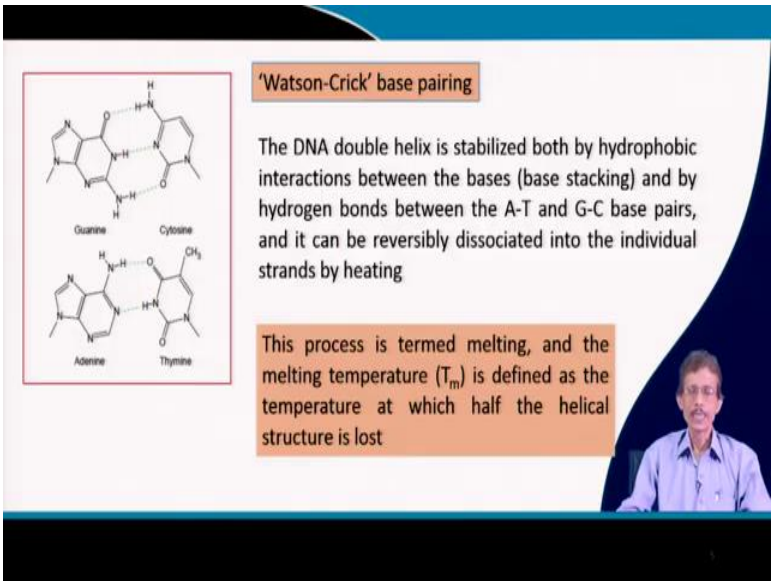
So, the overall base composition is varied widely and always we see that some amount of adenine, thymine, guanine and cytosines are the same amount of G and C that means A is balanced with T and G is balanced with C. That means they are coupled together, that means A will be hydrogen bonded to T and G will be hydrogen bonded to C also.

Then we have seen that is known that high quality X ray photographs of DNA fibers, since the fiber like of material are consistent with your helical structures composed of either two or three

poly nucleotide chain that means, you have the three now of the three dimensional structures and these hydrogen bonding interactions between these bases are there.

So, if the bases are there one after another you get a single strand and the another bases from the other strand you get it and if they are coming together for hydrogen bonding interactions, two hydrogen bonds or three hydrogen bonds you can have. So, three of these hydrogen bonding interactions so you get the corresponding helical structure.

(Refer Slide Time: 8:00)



So, you see now that G that means the guanine is attached to cytosine and adenine is attached to thymine. You see, they are, how nicely they are hydrogen bonded. That is why the hydrogen bonding interactions are so important. And we should always try to understand the capacity, the capability of these hydrogen bonding interactions, we know that one of the hydrogen bond is weak, but we have seen that in water molecule even in ice also the central water molecule is hydrogen bonded, four hydrogen bonds are there with surrounding water molecules.

So, is a nice packing in the crystal lattice also, you have good packing of water molecules in the ice lattice, they are stabilized by hydrogen bonding interactions only. So, this structure can also be stabilized and you can have the energy gain also through this particular hydrogen bonding interactions.

So, in one case is not always is very easy to get that particular pairing basically the guanine is paired with cytosine. So, you see that one particular type of base is giving you that particular complementarity. So, is the complimentary to cytosine, that is why guanine is attached to cytosine only, guanine is not attached to thymine.

Similarly, adenine is attached to thymine and the hydrogen bonding nature, hydrogen bonding type and hydrogen bonding structure is different, because the NH2 function is involved, NH function is involved and CO function is involved. That is why we are talking about in our previous class about the protein, the polypeptide chain also, the hydrogen bonding stabilization from the backbone of CO NH. So, these are the two parts only.

Here again you have the CO function as well as the NH functions. So, CO and NH functions are doing the miracle in terms of the stabilization of the different structures and those structures. are responsible for everything and those structures of responsible for your stabilization and also we will find how the coiling, the coil will now can have.

When you have these two bases, the purine and pyrimidine base pair, these two persons again Watson and, Professor Watson, Professor Crick, they considered that you can have the hydrogen bonding interactions, in one case three hydrogen bonds are formed, in another case two hydrogen bonds are formed and they are considered as a very important part and they consider it as the Watson Crick base pairing.

That gives rise to the double helix structures, that we call as the Watson Crick model. And you can have that not only the hydrogen bonding interactions, but also you can have the hydrophobic interactions between the bases, that is how the bases are getting stacked. So, when you have these interactions to hydrogen bonding, because these are the planner parts, the bases are planes, the guanine, cytosine, adenine and thymine these are the planner part. So, you can have one after another the stacking.

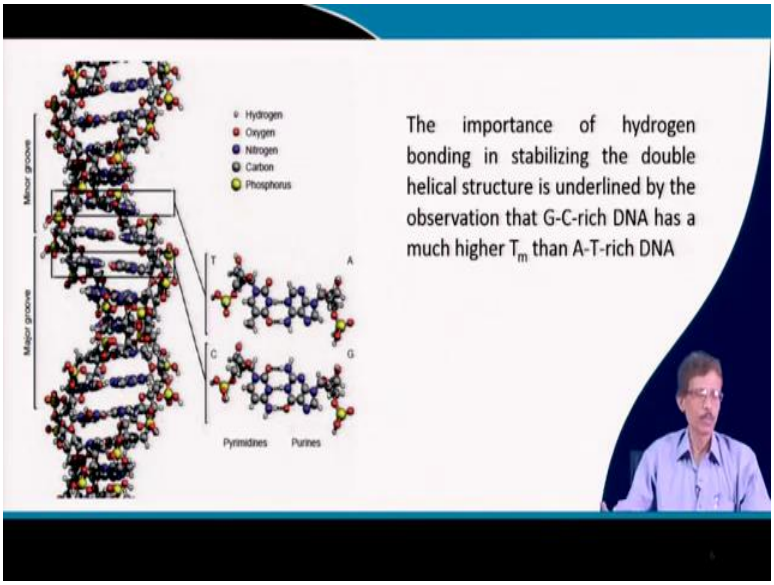
So, the pi pi interaction, the pi pi stacking can have also with you. So, the base stacking is also important as well as the hydrogen bonding between AT and GC and it can be reversibly dissociated in the individual strands by heating. So, simple temperature change basically can destroy the hydrogen bonding interactions and your chains can be released.

So, how the heating effect can be useful also. So, simple determination of that particular temperature where the melting process is taking place is also a characteristic information for all the different types of DNA molecules or the DNA double helix structure. So, basically this particular work basically, so all our Nobel Prize winning work.

So, initially the amount of efforts they have given to identify what sort of bases are there because no such access structures are available to them. So, they try to find out some particular very basic informations the way we do for our samples, whether you have a organic sample or a metal complex in your hand, we try to get the corresponding melting point.

Similarly, for DNA also, this melting temperature is the characteristic one where we get a value of that particular temperature and half of the helical structure is lost because is a melting process that means, the structure is not there. So, when 50 percent of this is lost, so there are mechanisms, there are instruments. So, those instruments can take care of measuring these Tm values that we consider for the polymer samples and all sometime we call as another Tm is the glass transition temperature. Well, this is Tm value, the melting temperature value is attached to DNA melting.

(Refer Slide Time: 13:06)



So, you see, do not worry about the very big structures, but it is a very nice structure as always we can have there is cover page of any book, anywhere you will find, even for your school days books, your CBSE books also you have seen that these helical structures and basically these are

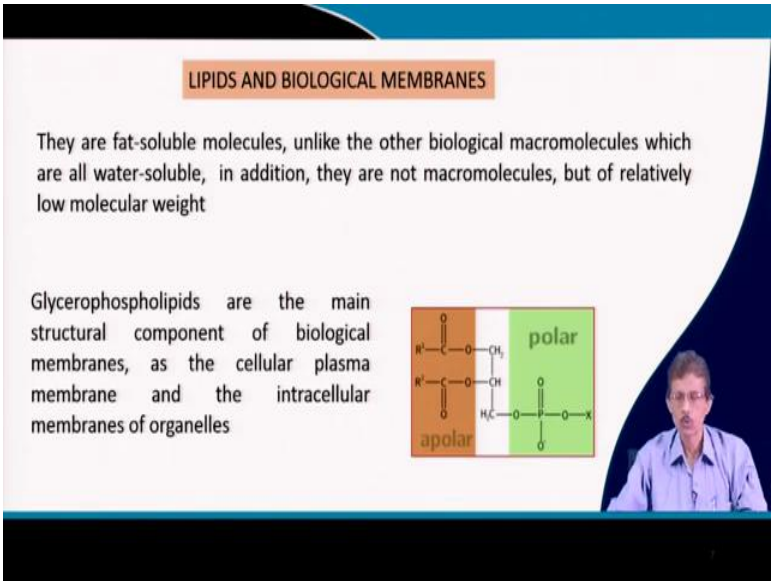
stairs. So, when you have the bases are paired up you have the stair, so stairs are there and this particular helical structures, the helical staircases are there. So, all these pairs are there starting from hydrogen to phosphorus you can have, so you can look at nicely where you have hydrogen and where you have phosphorus.

And this ATGC pairing. So, basically ATGC pairing is basically you can have, so, here you see that this particular point and these particular points so two points you can locate, this particular rectangular boxes. So, rectangular boxes are trying to locate basically two areas where you have the AT pairing and where you can have the GC pairing or CG pairing.

So, AT pairing and GC pairing can give rise to some space between these two strands. And those two strands are basically important for stabilizing the structures. And the importance basically is underlined by the observation that the G-C-rich DNA has a much higher Tm than the A-T-rich DNA. Why?

Because your GC can have three hydrogen bonds compared to your AT part. So, the number of hydrogen bonds will immediately give us that very important information that what magnitude of Tm you can think of, and if you find that particular value is directly related to the number of AT pair and number of GC pair sometimes we have to find out that for the whole DNA molecule.

(Refer Slide Time: 15:00)



Then we will just come to your other part of those biological molecules the lipids and the biological membranes, how we basically form the lipids and how the biological membranes are formed. So, what are lipids basically? So, we all know carbohydrate because the carbohydrate part, the sugar part already we considered, because our lives basically, we are a biological molecule itself also and we are trying to have good informations about, we will be talking about the metal ions.

So, whether your metal ion can interact with carbohydrate, whether your metal ion can interact with protein molecules, whether your metal ion can interact with the fat molecules, so protein, carbohydrates and fats, we are studying all these from our school days. Now, if you know that where you have the lipid and if the biological membranes are available, because we know that the metal ions can pass through certain membranes, and we can have certain membrane potential depending upon the concentration of the metal ions from one side to that of your other.

So, by textbook definition, they are basically the fat soluble molecules. So, the very basic thing always you try to remember, you understand is any time people can ask you, what is that textbook definition of lipid, what do you mean by biological membranes. So, these are fat soluble molecules and they are not like other biological molecules what we are discussing so far that your protein molecules, your carbohydrate molecules and sometimes your DNA and RNA molecules.

So, these water molecules are not macromolecules like those molecules are macromolecules, DNA, RNA and all, the proteins are also macromolecules, but are relatively low molecular weight molecules. So, very easy to understand, very easy to handle also. So, this low molecular weight lipid and biological membrane molecules we can have.

And again now, a mixture of these two, a polar part you can have and a non polar part also. So, if you have a non polar part and the polar part together, you can have some very good informations when you can have the corresponding folding, we call as micellization, the micelle formation, and like your soap molecules and all.

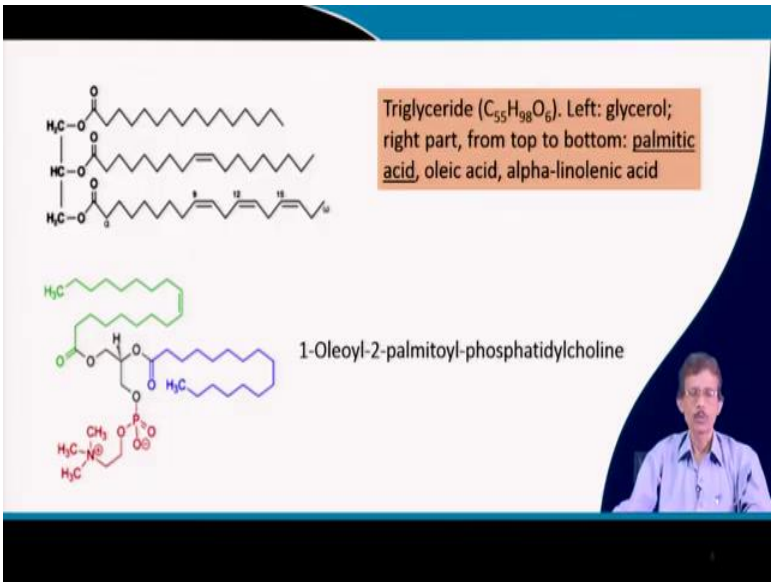
So, you have the polar ends that means, the charged ends and the nonpolar end or the apolar end that the uncharged ends. Then one such example of the molecule, why I am talking about everything together, that glycerophospholipids is not only a simple lipid, because these are one

part to the other, we all know what is the fat, what is oil and what is fat from your school days, you know.

Now, we bring something the poly alcohol, glycerol we bring and the glycerol based lipids we can have and that glycerol base lipids we bring some phosphate groups there also. So, it will be glycerophospholipids and which are mainly the basic component when you try to get the biological film or the biological membrane, where we can enclose the cell because the cellular plasma membrane and the intracellular membranes are different organelles.

So, you have to enclose those organelles by some film or by some membrane, there are phosphate groups. So, phosphate groups can interact with some other part and there are some Ester part, the typical Ester parts, the left hand part, the apolar part basically are the ester part. So, that ester part, so two are the typical esters bearing R1 and R2 group and another is the phosphate part.

(Refer Slide Time: 19:06)



So, simplest one is your triglyceride, we all know in our body also sometimes we go to doctors and doctors basically tell us about the amount of triglycerides in our (())(19:17). Body also they are in the veins and arteries also, they are getting deposited also in the different arteries and veins, but they have the typical role to play in our functions.

So, triglyceride the typical molecular weight. What is there in that triglyceride? We can have that left hand part is your glycerol unit CH2 OH, CH OH, CH2 OH, we all know that is the glycerol part. So, the glycerol part if it is attached to some Ester component to that of your long chain fatty acid we know that is your fat or oil.

So, if you have three different parts basically, one is your palmitic acid which is underlined here the top one, then oleic acid and alpha-linoleic acid, how they are different, we see that the palmitic acid is the saturated acid. Then one unsaturation and then three unsaturations at number 9, 12 and 15 and basically the positions also is defined as the omega 3 type of, omega 3 unsaturated, omega 6 unsaturated, omega 5 unsaturated, we know that the unsaturated acids are very useful.

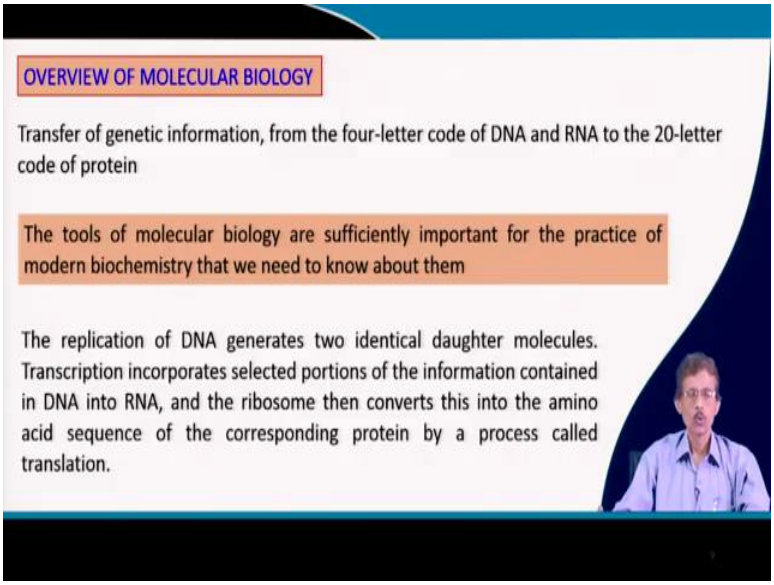
Then another one not only that triglyceride part or the fatty acid part, we also attached to it is the Choline part, many of these molecules will have see, we will see also that there are many such groups available where you can have the Choline part. So, what is that Choline part?

So, Choline attachment whether it is a charged or a neutral one that we can also see. Now, we can find that already we have seen what is oleic acid, oleic acid we will have one double bond only. So, this part basically the top green part we find that is your due to that of your oleic acid, and this black backbone is nothing but your corresponding glycerol part.

And this part basically the palmitic part, so the palmitic acid part already know that is the fully saturated part. So, these two are same, so what we can have the oleic acid part is there, the palmitic acid part is there, but instead of alpha-linoleic acid, you can have a phosphatidylcholine part. So, this is the choline part. So, tertiary nitrogen 3N amine 3 part.

So, at the end is not NH2 function, not also NH3 plus function, but is N any three plus function, which is basically known as the corresponding Choline function. So, the Choline function sometimes is very important and you find that the phosphate can have a negative charge the Po minus, this nitrogen has a positive charge, so is basically (())(22:07) ionic part and which is also very important on this particular lipid layer. So, they can have some very important reactions to perform.

(Refer Slide Time: 22:19)



So, these are basically what we are talking about the corresponding reaction vessels or sometime we find that metal ions can go, the MIs can go to interact with those particular species like DNA, those particular environment, so how the environment you can have, so because all these reactions the reaction dynamics, the reaction kinetics and the reaction thermodynamics also.

How we can understand the potential at which the electron transfer is taking place, at what pH the proton transfer can take place and at what particular rate basically the rate of electron transfer is all will be dependent on the corresponding reaction environment, the reaction chamber or reaction vessel, the way we call it as the vesicle formation also in the biological world also we find that there are some vesicle formation.

So, basically in terms of your molecular biology informations and genetic informations how we can transfer. So, we have a four letter code for DNA and RNA synthesis. So, the synthesis is also a very typical one. Then we find that what is that particular code you have to have the code because that on the code basically, we will be getting the corresponding 20 letter code for protein synthesis is basically a genetic code material we call nowadays in the computer science, when you talk about computer science, what you are doing you are writing codes.

So, the term basically is well known, the nature is doing for us basically, the nature is writing that particular code that what particular DNA and what particular RNA will be synthesized. And

these RNA and DNA will basically dictate therefore, a more higher number of letters involved code which is your protein code for 20 letters for the protein synthesis.

So, these molecules basically can dictate that which particular protein will be synthesized, forget about the metal of protein structures. So, the synthesis of the protein is important, first DNA and RNA molecules will be involved over there for that particular synthesis. And then some of these proteins are having useful donor groups and those donor groups are there to trap the metal ion to bind the metal ions.

So, the techniques or the instrumentations or the methodologies which are available for Molecular Biology understanding which is also very important in understanding the modern day biochemistry that we need to know out of these because little bit of molecular biology we have to know it otherwise it would be not very easy to understand everything, because once you know about the DNA structure, we talk about the DNA replication.

So, that particular replication of DNA generates two identical daughter molecules. And then we have to go for the transcription process and which basically nothing but the incorporation of the selected ports and basically one part of the information content in DNA into RNA. So, the information goes from DNA to RNA and the ribosome then convert this into the amino acid sequence of the corresponding protein by a process called translation.

So, these are nothing but simple definitions all see. Because, once you see the typical definitions, you will be able to understand many things very quickly and very nicely that why we should study DNA when we are talking about metal ions and all these things, but now, if I say that we can characterize the DNA molecule or the RNA molecule by knowing the Tm values, so these Tm values are important.

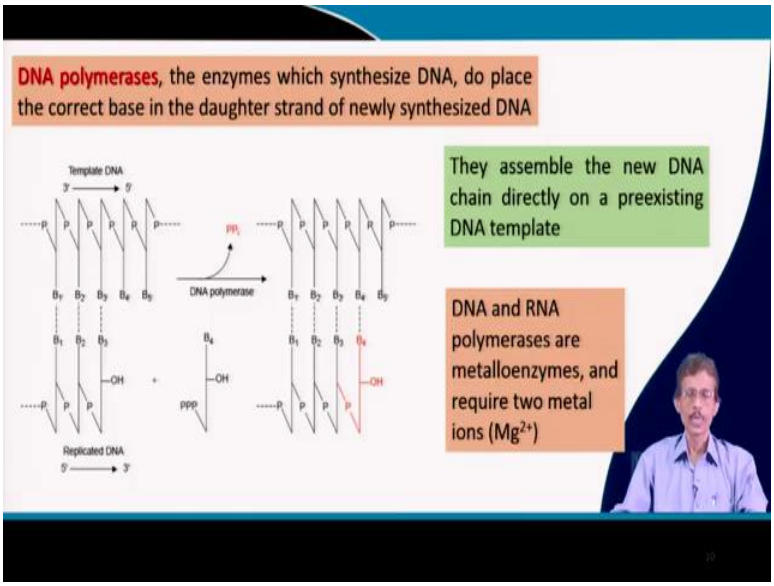
Now, if we see that you have the nitrogen donor bases and those nitrogen donor bases can be available for interacting with the metal ion, as we have seen that our body or any living organisms, the plants, the other animals, they have huge amount of metal ions, the storehouse of the metal ions are there.

So, if the bases are there, will it be possible for us to keep them separate? No. So, if your Rn is available, and if your nitrogen base donors are available, so can nitrogen will come and interact

with the iron. So, these chemical interactions are always have some meaningful role to play. Either you have to compartmentalize these, you have to discard those molecules, you should not allow them these molecules, these ligands basically.

In my definition, what I can say is that from that particular definition, you have to segregate those molecules such that they should not come and reach the metal ions, such that those very good ligand molecules are coming and interacting with the metal ions. So, during this particular process the replication to translation, you have to have the corresponding molecules involved, which are responsible for synthesizing the protein molecules.

(Refer Slide Time: 27:42)



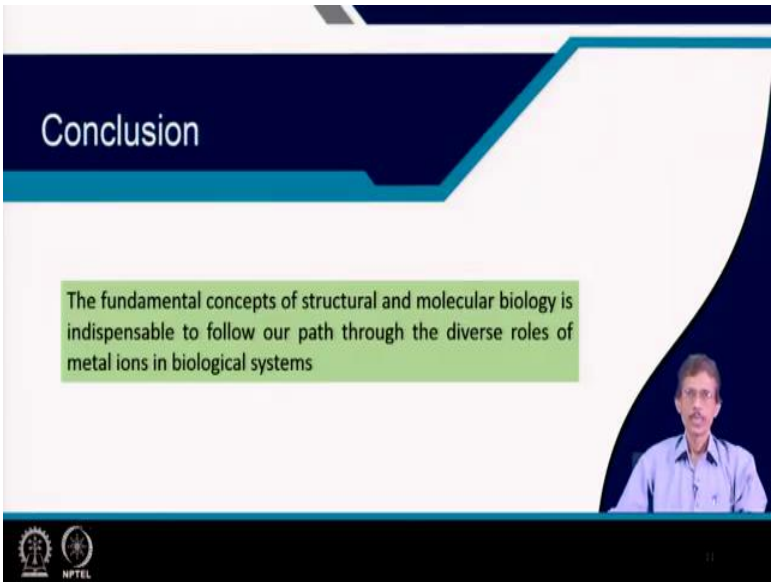
Then, one important molecule is your DNA polymerases because, their metal ion is involved. So, is a very big one (())(27:49) cartoon diagram, again the cartoon formation that B says Br phosphatase and the black straight lines are your backbone. So, when you have the DNA polymerase is interacting, you have to release the corresponding inorganic phosphates PPIs, the diphosphates basically.

So, they basically go for the correct placement of the bases in the daughter strand in the newly synthesized DNA. So, DNA synthesize are basically nothing but new DNA chain will be formed directly on the preexisting DNA template. So, this is your template and on it you are putting something else such that you get that.

So, many such DNA and RNA polymerases are nothing but metalloenzymes and they require two metal ions. Now, you see that the metal ion is not iron, not copper, not zinc. So, if you start studying bioorganic chemistry or the biological inorganic chemistry from this point basically, I simply jump on the magnesium coordinates in chemistry which we do not know much.

Magnesium we know as the magnesium present in organometallic world, in organometallic chemistry in many other main group chemistry and in the solid state chemistry, but magnesium in biology, which is a very important part people are nowadays people are start learning nicely about this, but the various important thing and very important information is what you can have where you find magnesium in biology. So, you will find it in the DNA polymerases. So, this magnesium definitely will be requiring there for interacting with phosphate molecules and they are important in this important molecule.

(Refer Slide Time: 29:39)



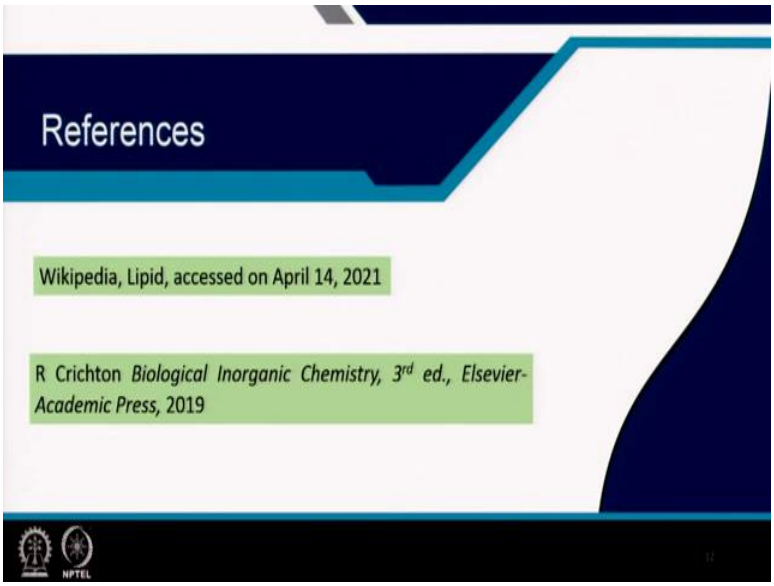
So, what we can have seen in this particular class is that the some very basic concepts not the fundamental even also. The basic concepts what we can have for the structural part and the molecular level part of the biological world because we have to follow these where the roles of the metal ions in the biological systems we should be able to understand.

So, definitely these informations are indispensable, you cannot discard it, you have to read it, you have to study it to follow our path, for the diverse roles of the metal ions in the biological system starting from the metal ion homeostasis, the metal ion assimilation process. That means,

how you synthesize hemoglobin and myoglobin, whatever we are talking, we will be talking mostly about the structure and its functions.

But if we go beyond that, if you are really interested, you will be interested to know how the hemoglobin and myoglobin is forming from the scratch, from the very basic building blocks, how the hemoglobin and myoglobin molecules are forming, how our body is producing, the way it is functioning we know now, where it is malfunctioning we know that is why you go for taking medicine, we take the drugs and all these, how it is malfunctioning, but how it is synthesizing in our body that we do not know much.

(Refer Slide Time: 31:15)



So, the books basically what we will cover again from the Wikipedia page, the lipid part you just read it nicely is only one sometimes single page with having 10, 20 or 30 references. If you have time, you just go through those references also. And then one very important book, mostly this book is very, is basically for this particular class is the Bible for me. Because mostly we will be following this particular book is the biological inorganic chemistry in the most latest one is the third edition. So, the version what we can have is also the latest versions, the latest informations by Robert Crichton. Thank you very much.