Inorganic Chemistry of Life Principles & Properties

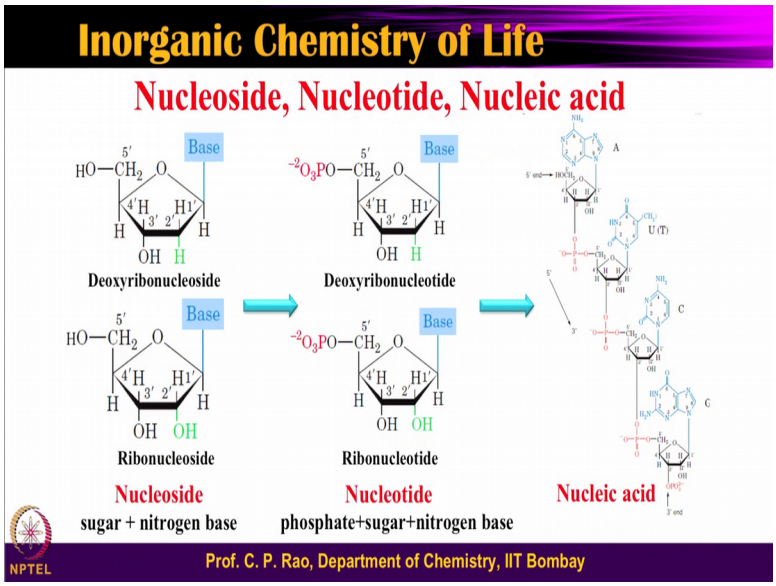
Prof. C. P. Rao
Department of Chemistry
Indian Institute of Technology, Bombay

Lecture - 07
Nucleoside, nucleotide and nucleic acids & DNA: An introduction

Good afternoon, welcome you back to this course on Inorganic Chemistry of Life principles and perspectives. So, in the previous lecture we have looked at some aspects 11 to the protein, starting from amino acid to the peptide to the polypeptide, where in we can see N terminal, C terminal directions then followed by the secondary structures which are nothing but alpha helix and beta sheet structures. And then when you have all such structures a built in the entire polypeptide chain, what you have is a tertiary structure where you get there overall topology of the protein structure and such tertiary structures joined together in this particular kind of a symmetry, to give a quaternary structure and then quaternary structure is a perfect you know structural position for a protein to give its function.

Even some cases tertiary structure can also be useful as a as a functionary state. Wherever you have one sub unit and when you have more than 1 sub unit you have a quaternary structure is formed. So, this is what we learnt in the previous class and in this class, let us try to look at other kinds of biomolecules, the biomolecules suggest the nucleic acids and you know in the nucleic acid terminology is bit different from that of the terminology or the proteins.

(Refer Slide Time: 01:53)



In the proteins case you have a peptide and you have an amino acid. So, amino acid turns to protein, here it is other way round, it is a nucleotide which turns to nucleic acid. So, you have nucleoside, nucleotide and nucleic acid as for the hierarchy is concerned.

As you can see from this particular slide, that you have a nucleoside; nucleosdide has the ribose and the base and that is a nucleoside. And nucleotide has the phosphate the ribose and the nucleic base and that is the nucleotide. The one which is shown over here has no oxygen in the second position. So, since there is no oxygen is a second position it is called Deoxy. Therefore, diseases called deoxy ribonucleoside this is called deoxyribonucleoside and this side you have the additional unit is the phosphate. So, therefore, once you have the phosphate and sugar the nucleic base together is nucleotide.

So, this is called deoxynucleotide, I will come just in a while the difference. The once which you have the two structures the below, these 2 structures are very similar to those are the 2 on the top, except for in the second position you have a OH; that means, oxo. So, this is referred as a ribonucleoside; so ribos plus the base ribonucleoside, the phosphate ribose plus the base ribonucleotide. So, ribonucleoside and ribonucleotide and. So, the difference between these 2 is the second position. In case of DNA it is Deoxy Ribonucleotide and RNA Ribonucleotide. Now and deoxy ribonucleic acid and ribo nucleic acid these are the 2 different types and in fact, the you will come to know it will

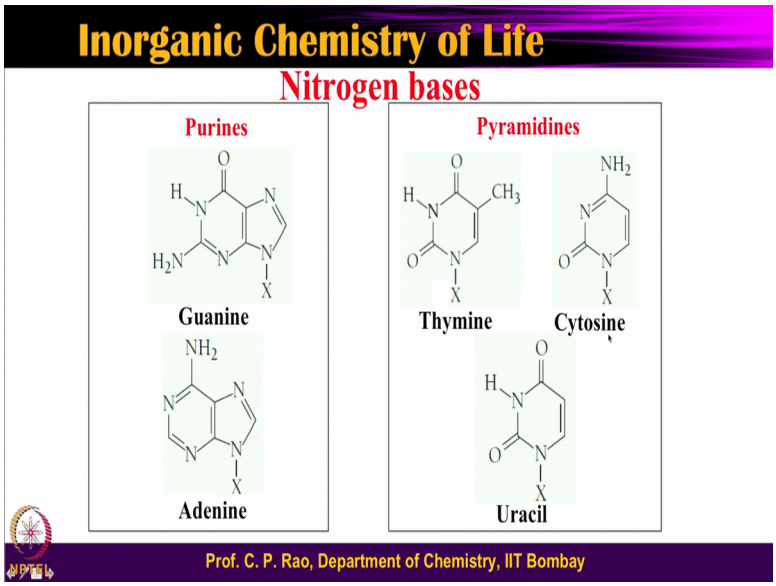
much later stage when we talk about the enzyme ribonucleotide reductase, I will be bringing more details but let me bring 1 or 2 differences important points in this context.

The important points in this context are in its second position, the O OH there have a nucleoside, these are the inputs to the human system whereas, deoxy the component of the DNA is not an input of the body. So, what we the input of the ribosill nucleoside nucleotide can be converted into deoxy so; that means, you essentially you remove one of the oxygen, then it will go to this. So, it is called deoxy. So, there are enzyme which does that. So, deoxy ribonucleotides are required to make the deoxy ribonucleic acid and therefore, the deoxy ribonucleotide is noted entry point to the human system and this is made out of this ribonucleoside by deoxygenation process on this ok.

So, this is one thing you keep in mind. So, when we come to the enzymes. So, deoxy ribonucleotides we will explain all this details this. On the right end what you have is if you connect each of this nucleotide if you connect the ribonucleotides you will get a RNA, if you connect deoxy ribonucleotides you will get DNA. So, so you have the nucleic base, the sugar, the phosphate, the sugar, the nucleic base, the phosphate, the sugar the nucleic base and the phosphate. So, what is the back bone in this? The back bone in this is the sugar, phosphate, sugar, phosphate, sugar, phosphate what is the backbone in the protein? The protein backbone is C alpha CO carbon and N, C alpha CO N C alpha CO N.

So, that is the backbone and what is the side chain or group. So, in case of nucleic DNA, it is the base nucleic base whether it is RNA whether it is DNA it is the nucleic base, which is the kind of a equivalent to side chain. We do not call this as a side chain in DNA terminology we call it is a nucleic base so, but main chain is the sugar phosphate, sugar phosphate, sugar phosphate ok. So, you have by virtue of having the ribonuclease, you will have a RNA by virtue of having the deoxyribonuclease you will have DNA.

(Refer Slide Time: 06:55)



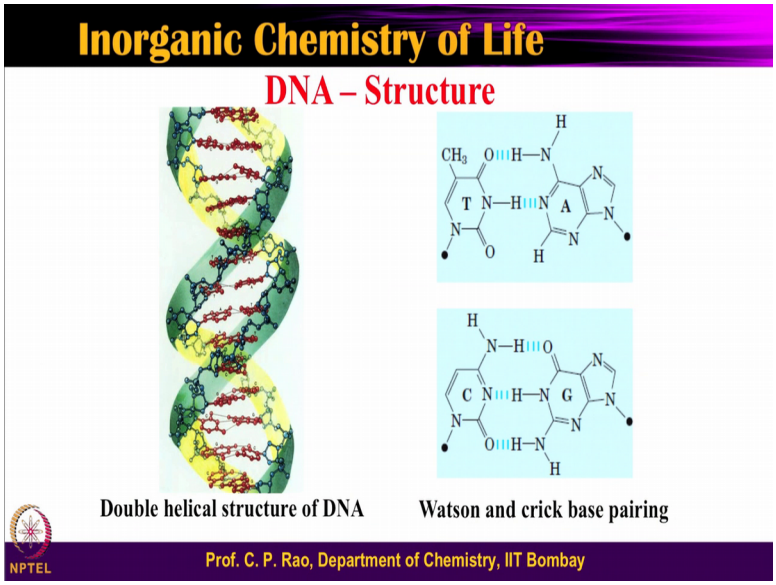
So, let us look at the nucleic bases in little better way, the nucleic bases or not like the you know base like hydroxide or carbo needer or something, these bases are coming from nitrogen containing centers.

So, in the nitrogen having a load pair is a base. So, what kind of base is this? Ok this is Lewis acid Lewis base concept Lewis base. So, I am sure I will be explaining very soon about the Lewis acid Lewis base those concepts relevant to all these things. So, right now take this nitrogen centers having the base. So, there are 2 types of nitrogen bases, one is called the purine and other is called the pyramidine how do you remember? If there are 2 rings you can take it as a purine, if it is one ring only it is you take it as a pyramidine. So, purine and pyramidine what are the purines? Purines are guanine and adenine; so guanine and adenine whereas, pyramidines are - thymine, cytosine and uracil ok.

So, these are the five nitrogen bases or in another words nucleic bases, which present in the RNA and DNA. RNA DNA or more or less some excepts for in the second position the ribos, in one case you have a OH other case you have only H. So, the one which you have H is called deoxy so, DNA thing. Now you got in the nucleic acids, backbone is a sugar phosphate, sugar phosphate. So, the phosphate is connected and both the directions phosphate you know is corrected at both ends. So, therefore, you have. So, phosphate a sugar in one side and phosphate in between and again sugar on the other side. So, its a kind of a bridging. So, phosphate bridges in these things ok. So, now, having seen the

nucleoside and nucleotide and nucleic acid, and then we have seen the side chain here and nothing but nucleic bases.

(Refer Slide Time: 09:02)



So, if one more do expect some structural behavioral like that in proteins, you will find somewhat similar kind of things. So, these are again helical kind of a structures that you have.

So, the helical structure that you what you have here would basically talk about the backbone. Backbone is the nothing but the phosphate sugar phosphate sugar, and onto that anchor or the basis you see this is red ones are the basis. So, this is one strand and this is the other strand. So, there are 2 strands are there and in between there are basis of their. So, what is that kind of, what is that stabilizing such kind of structures. So, since there are 2 strands and the 2 helices the structure is called double helical structure. So, you must have known even very early stage in your high school with the DNA has a double helical structure, and that is because of these ones. Now how these structures stabilized? So, to understand that look at on the right side on the top for example, you have the 2 nucleic basis that you nucleic basis or interacting through hydrogen bonds N H O then N H N.

So, you have the N H N hydrogen bond and N H O hydrogen bond. So, what are the partners here? The one of the partners thymine other is adenine. So, it is an AT pair or TA pair it does not matter the one and the same. So, you can see that. So, in one context this

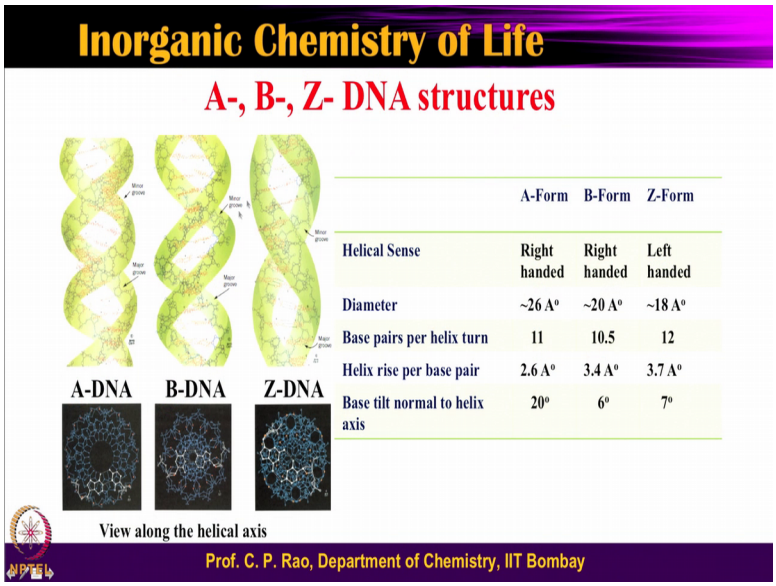
is a donor for hydrogen another context is the accepter. Similarly exactly in mirror this acceptor here and donor here; so donor and accepter here we refer to the hydrogen hydrogen donor hydrogen accepter. So, the total interaction is referred as the hydrogen bond formation, you know that such hydrogen bonds which have formed would have a good amount of stabilization energy in the order of 1 to 4 or 5 kilo calories per mole.

The same thing is true even in the protein structure, when I talk to you about alpha helical structures or beta sheet structures again you have the similar kind of. But the hydrogen bonds they have or N H and O, but here we have NHO we also have NHN. So, as the electro negativity decreases, the strength in a bond also bit decreases. So, you have AT pair this is called which famously known as AT pair so, also called complementary pair.

And see the structure below there are again 2 there are nucleic bases are there and one is called the cytosine other is called the guanine. So, a guanine cytosine or not interacting just with the 2 hydrogen bonds like you have here, but rather their binding through three hydrogen bonds.

So, one is NH O, other is NHN, another is again NH O. So, you have very too strong and one somewhat will less strong hydrogen bond whereas, here one strong one less strong I am not saying weak, but somewhat will less strong ok. So, therefore, over all the interactions are very good to stabilize. And now if you take this concept and see here between the 2 nucleic bases, these are basically stabilized by such kind of hydrogen bonds. We cannot see here that is why expanded in show. So, AT is one pair GC is another pair. So, these kind of a pairing is referred you click base pairing is called a Watson crick base pairing and this is the reason why these structure the DNA is stabilized in its double helical fashion ok. Now it is not only one kind of as a double helical structure, there little differences for one double helical to another double helical to another double helical.

(Refer Slide Time: 13:06)



	A-Form	B-Form	Z-Form
Helical Sense	Right handed	Right handed	Left handed
Diameter	~26 A°	~20 A°	~18 A°
Base pairs per helix turn	11	10.5	12
Helix rise per base pair	2.6 A°	3.4 A°	3.7 A°
Base tilt normal to helix axis	20°	6°	7°

You can see somethings are clear this is little bit more elongated as compare to this and this is much more elongated than these 2.

So; that means, there are some structural changes though all of them are double helical structures ok. These structures are referred as A B and Z ABZ, DNA structures; A-DNA structure, B-DNA structure, Z-DNA structure ok. So, there are some parameters given over here before coming to the parameters let us look at this pictures given just below each of these what are these pictures you think? Suppose if I take an access here and if I put my eye here and look down, then I can see this whole thing crumpled 1 over the other and there is a interior opening and that is what the interior opening is and all these helices will fall over the other it will become like a circle here.

So, it this is the whole circle and the interior. So, that is what you have the entire thing. So, this comes you look from the top. If you look from the top you will find this kind of structure. So, similarly if you look from the top from the B-DNA, you will find a structure of this type, and if you see this one it will find structure of this type. So as you can see the interior the portion which is quite hollow in A stack of structure is much less hollowness, you can see and much more compact. So, somewhat lesser filled somewhat more filled somewhat much more filled: A time DNA, B time DNA, Z time DNA. So, there are three times; now let us look at the on the right side what we have.

So, the A type and B type or the right hand in helices and the C this Z type is the left hand helix. So, as you can see it is going like this whereas, here it is going like that. So, this is called right hand and left hand, I am sure you are aware right hand is going in the clock direction clockwise direction and the left hand is going towards the anti clockwise direction ok. So, what are the difference is that you can? You can see you can see the differences in the diameter. So, the diameter for a is 26 angstrom, that is 2.6 nanometer, 20 angstrom 2.0 nanometer and 18 angstrom 1.8 nanometer. So that means, you are the cylindrical portion that diameter is reducing, I am going from A to B to Z time almost by 1 nanometer or about 8 angstrom unit.

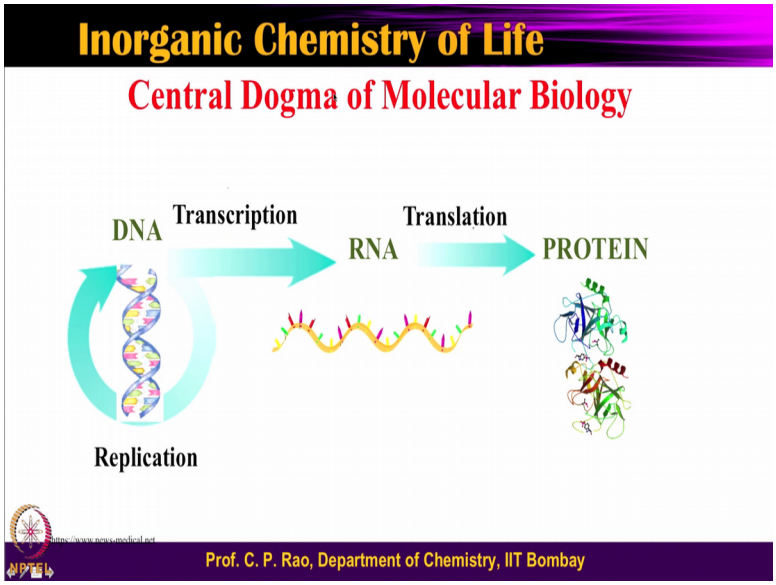
That is quite a substantially different, that is why you see the structures look very different when you look down the central axis. Then look at other aspect which is not of course, so clear from here, but you can take it is granted, this B space how many helix. So, you know helix. So, from here to here was 1 screw and this base pair for a helix turn there are 11 base pairs are there, 10.5 base pairs -are there 12 base pairs are there ok. So, this how the 1 is this way, one is bit elongated other one is much more elongated kind of structure and now so for each turn how much is the height increased; the increase in the height which is called helix rise for base pair. So, in stuff for turn for each base pair, its divided by the number of turns a number of base pairs. So, we will get for the each base pair.

So, its 2.6 angstrom in case of a A structure and 3.4 angstrom in case of B structure and it is about 3.7 angstrom in case of Z structure that is easily understandable this got so, much elongated. This is narrow, this much wider. This is more elongated this is a perfect kind of a almost symmetric type you can see so symmetric very nicely. Now look at their other parameters there is the base pairs or in each of this will not have same kind of orientation. So, if you take 1 base pair and the complementary, how well they are arranged if there exactly head on the angle will be 0. So, there is that will be 0 with respect to the normal with respect to the access that you have. Now if you look at these ones A case, the base has a tilt normal to the helix is about 20 degrees and this has only 6 degrees and this has 7 degrees so almost very similar, but we between these 2.

So, these are some of the differences that comes for these three kind of a structures and need to tell you that this A B and Z kind of a structures or inter convertible by changing certain conditions, condition such as salt concentration in which these are immersed and

the humidity or water content humidity content etcetera, so because the humidity belongs. So, compete with the hydrogen bonds that you have and the salt concentration will put the kind of a gradient. So, you under the gradient and under the water dipoles so, you can make its structure change from A type of a DNA to B of DNA to Z type of DNA structures ok.

(Refer Slide Time: 18:47)



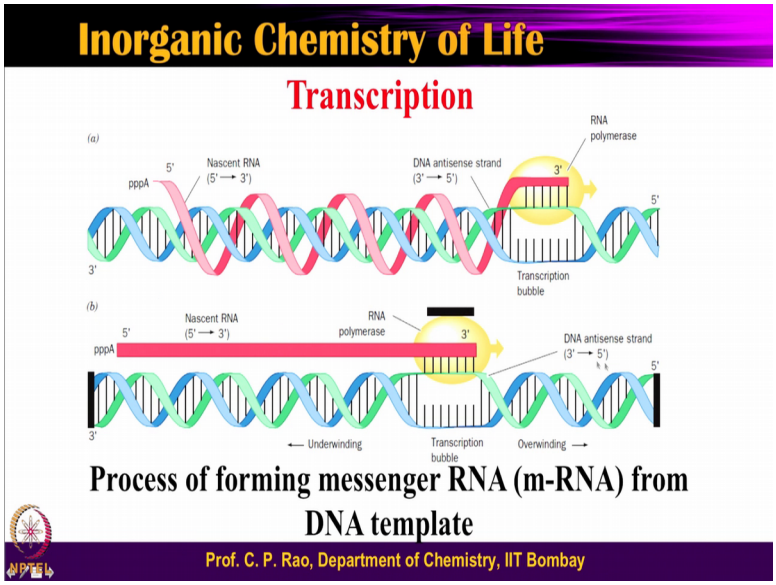
So, there is something very basic on the DNA ribonucleoside nucleotide and nucleic acid, and then DNA A type, B type and Z type. So, I think that basic information is more or less good enough for this particular course though a lot more biological importance and essential aspects are there which are not a part of this particular course at all ok. So, having looked at we have looked at the protein structures, the protein how it is formed the nucleic acids, how it is from nucleic acids structures now will let us smoothly get into a connectivity. So, how this proteins for example, synthesized? So, how are the proteins synthesized and to this you can just have a look at this is in any other molecular biology book even a any other biochemistry book having a chapter on the protein synthesis on molecular biology aspects.

You would see a title called central dogma of molecular biology. So, is the central to the molecular biology? So, what is central to the molecular biology? So, you have a DNA, you could transcribe. So, it is a transcription process and transcription will give a transcripted version and transcripted version is used to express and expressing is called

translation and this translation will give protein. So, DNA to RNA to protein is generally referred as the central dogma of molecular biology. If you open any other molecular biology books biochemistry books you will never miss to see this particular thing ok. So, m is equal starting from DNA and what you need? You need a copy of this because copy of this and that copy is refereed as the RNA, which is nothing but m RNA I will give more details in the later slides this m RNA is used as a template to make to translate to translate the into amino acids and the amino acids are joined together to form the protein.

So, how were this translated? I am sure many of you might know that there is something called triplet code, three base pairs in a sequence is together will code for 1 amino acid. So, if there is some reading mechanism, it reading mechanism will read not individual ones as a triplet code so, that that triplet code will lead to 1 particular amino acid ok; so that amino acid. And the next amino acid in the next amino acid and each of these adjacent amino acids are being connected by the peptide bond and the protein is generated. So, you going from a DNA of this kind of a structure double helical to a protein may have a helix, may have a beta sheets, but not a double helix is a different kind of structure ok. So, one way you can say the protein is a kind of a daughter species coming from the DNA ok.

(Refer Slide Time: 22:09)



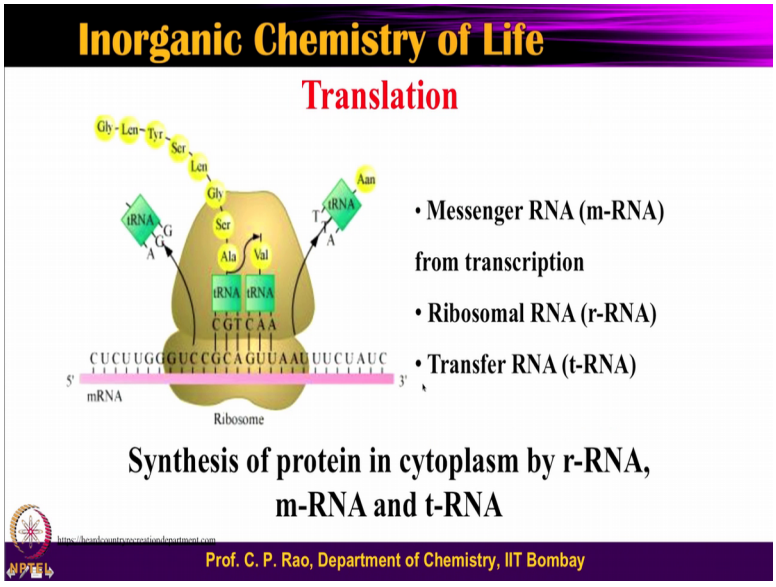
Let us get a little bit more detail at what immersion in the previous slide, what I said transcription. To get the transcription to transcribe the DNA so, you need to read the

DNA, and then once you read the DNA make a copy of that so is base essentially copying the DNA. So, to copy the DNA, DNA being double helix you need to unwinding ok.

So, therefore, you can see this is unwinding process this ok. So, unbound protein will be read and the copy is being made. So, as you can see over here. So, process of forming messenger RNA. So, there is copy is the called the messenger RNA and that is what is made out of this particular. And these are some details like from which direction 5 prime to three prime etcetera, which I have not introduced to you much. So, you do not need to worry.

Just like N terminal to C terminal, C terminal to N terminal input hence here you have 5 prime to 3 prime, 3 prime to 5 prime where is this is 5 prime and 3 prime is coming; is coming from a sugar moiety; so in the sugar you have a 1 prime 2 prime 3 prime 4 prime and 5 prime. We have seen if there is some change in the 2 prime form O H to H we call it is a deoxy just like that. So, similarly you need to use that. So, then you have a DNA and a sensing strand, then there is will open up then read and make a copy of this.

(Refer Slide Time: 23:58)



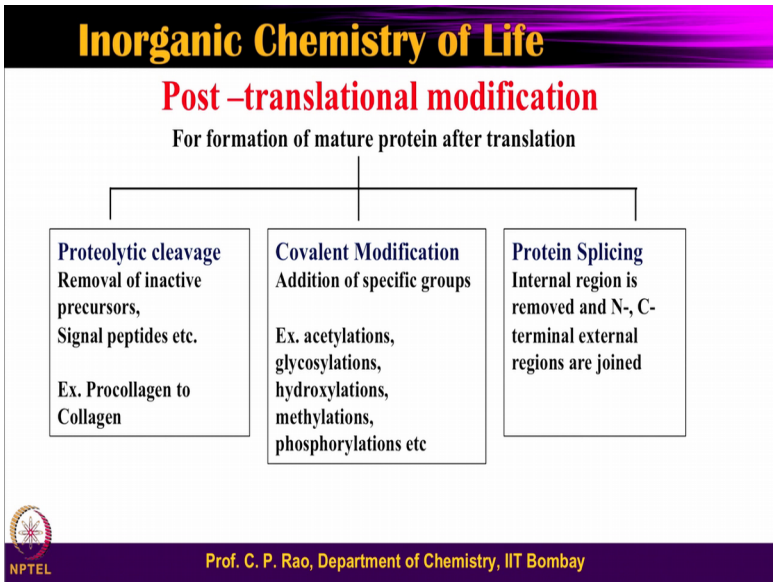
So, the making the copy of this is referred as the transcription process and this copy is used as m RNA this m RNA is referred. This is the copy which we got from the previous slide and this particular m RNA copy messenger RNA m stands for messenger.

So, basically it gives the between the DNA and going to the protein it is a messenger like system ok. So, this has got you can see very clearly C U etcetera etcetera and these are all nucleic bases and as I said three nucleic bases together will become it triplet code. And now this is entire process occurs at the ribosomal system and you can see I am not going in to more details in the ribosomal fractions. There are two different fractions the ribosomal, but I am not going into the details there. So, all that you need is to read a triplet code and this triplet code is transmitted to the t RNA and the corresponding t RNA is activated and each t-RNA will pick up 1 amino acid only.

That means there are different t-RNA is for different amino acids. So, the t-RNA basically makes a complementary contacts with the amino acid and brings it to this particular region and therefore, and drops it here. Then the next one the next one and then there is a other enzymes which will basically make a bonding between the 2 adjacent the amino acids, and as a result is not making the and building the protein. So, this peptide is built. So, this is the first one second one like that. So, keep shifting. So, you will shift here also. So, this is the track on which the ribosomal shifts and reads ok. So, ribosomal system is a mechanism to read translate to t-RNA and then hold it for bringing the right amino acid, then the right amino acid is being connected to the previously existing amino acid to form a peptide bond amino acid.

So, a messenger RNA m RNA from transcription then ribosomal RNA is and then transfer RNA and transfer RNA is specific for each one of the amino acid. So, synthesis of protein in cytoplasm by ribosomal RNA, messenger RNA, and t-RNA I think there is lot more details which I have not given, which biochemist and biologist study, but I have kept purposefully aside because that will be out of focus of this particular course, but knowing this basic thing is rather good enough. And after this protein is being formed actually many things happen and whatever happens after the protein is formed is called post translational modification.

(Refer Slide Time: 26:53)



You know translation means protein synthesis. Protein synthesis at the ribosomal is referred as the translational phenomena and post translational after being synthesized then lot of modifications. One of it is in fact, getting this structure protein folding is a huge research area protein folding and protein where it initially synthesize does not fold.

So, starts folding only when it acquires a minimum size of the protein etcetera. So, that protein folds only when it folds into the secondary and then tertiary, then we can see the functions as I told earlier. Besides this other things give a also happen. So, there are some proteolytic cleavages, some inactive precursors etcetera; these will be and signaling that means, starting etcetera those things will be removed after the total synthesis.

Then there are some kind of covalent modifications are also occurring such modifications if could be glycosylations, could be acetylations, could be hydroxylations, could be methylations could be phosphorylations. So, all of these when you see in a protein, they are not during the protein synthesis, but after the protein synthesis and protein splicing. So, internal region is removed and N and C terminals a regions are rejoined. So, by this time you have really made whatever the next level of the protein, which is suitable for utility for the rest the future purposes of this.

So, in other words at this stage you can expect a matured protein or a protein which can exhibited secondary tertiary and quaternary structures are being formed, hence this can be used for functional activity of this. So, in a fact in this particular lecture what I have

talked to you is as follows. Initially I have introduced in the previous lecture the amino acids peptide protein protein structures, then I have taken you into a new newer type of biomolecular systems.

In the newer type of biomolecular systems I have explained you the nucleic base nucleotide nucleoside nucleic acid and the double helical structures the; a structure b type of DNA, and z type of DNA and their characteristics. And these of being can be this can be modify or maintained or transformed by using a salt concentration variation or the humidity variations etcetera of all these things.

Then we enter into (Refer Time: 29:35) of how such a protein is being synthesized from DNA, DNA to m RNA to protein which is nothing but the central dogma of molecular biology and in this you make a DNA a copy of the DNA which is called the mRNA. The mRNA is used by the ribosomal system to use it as a template and translate into the amino acids then followed by stitching them together to form a peptide bond and to form a protein and then followed by that protein maturation and protein post translational modifications.

Thank you very much.