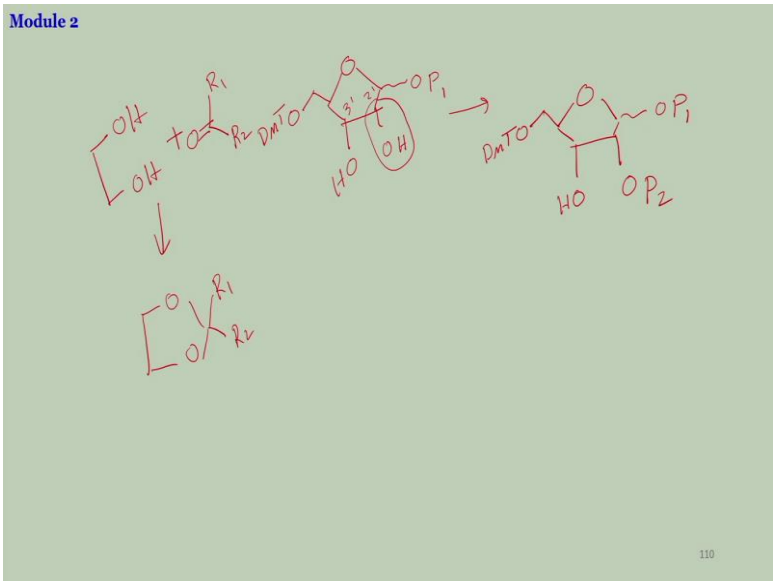
Essentials of Biomolecules:
Nucleic Acids, Peptides and Carbohydrates
Dr. Lal Mohan Kundu
Department of Chemistry
Indian Institute of Technology dash Guwahati

Lecture 09
Solid phase DNA synthesis

Hello everybody and Welcome back! So we are discussing about the synthesis of the nucleosides. In my previous lecture I have talked about how the nucleobases both purines and pyrimidines can be attached with the sugar moiety both in ribose sugar as well as the deoxyribose sugar. And then, we had, we have been actually discussing a little bit of sugar chemistry itself. How you can do the protection and deep reduction chemistry on the sugars on various positions of the sugars.

(Refer Slide Time: 01:07)



And we have seen in that if you have the ribose sugar then the C 1 prime position you can pretty much protect with any protection group separately because that that has very different reactivity. And then we have also protected the 5 prime with one of the example is with DMT that is the treat al, that is the tri phenyl chloride. And now we have the 3 prime H and two prime H. Now the question is how can you selectively rotate a two prime or 3 prime?

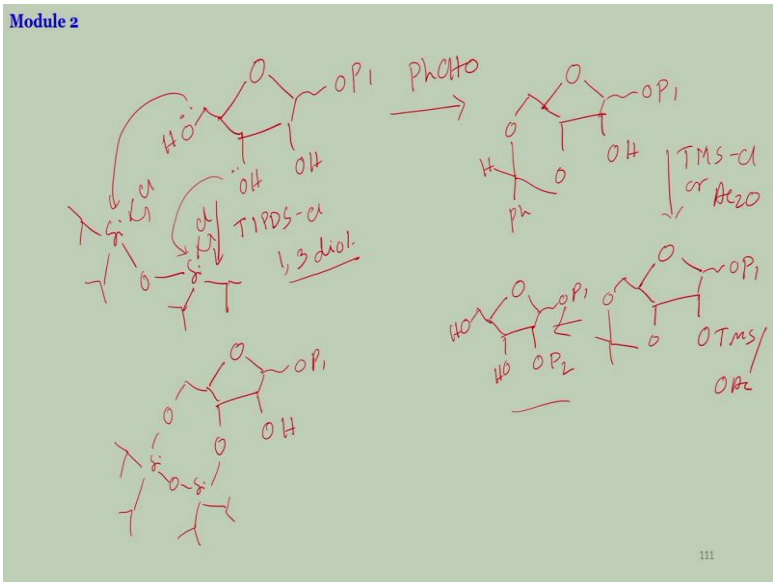
Now, if you want to protect the two prime position, so first thing is two prime and 3 prime these two positions they are reactivity are more or less same because both are secondary kind of secondary alcohols, now, secondary hydroxyl groups are present. Now, if you have remember there are few ways if we want to do the protection selective protection of the two prime position

there are few ways.

So our target is now, this is protect I am writing P 1, this is P 1 for one kind of protection. This I am keeping DMT on. So our target is protecting this P2 and keeping this free so that is our target molecule. Now so I will quickly go through it and then we can we will discuss about the nucleic acid synthesis the final part of it of this module. So there are a few ways to do this. One of these is of it.

If you remember that, if you have a dial and you treat this with a carbonyl compound or aldehyde or acetone then what they do is they form acetyl or Keitel this can R1, can be R2. It can be aldehyde or can be ketone. so this is called acetyl or Keitel so the same thing can be applied in here in this case the idea is that first you pro protect the other parts and then you can protect the C2 position.

(Refer Slide Time: 03:46)



So at the beginning without the DMT starting point is you have P 1 this OH and a free OH. Now, if you treat these this kind of old method actually with benzaldehyde then benzaldehyde also forms acetal formation but more preferably it forms the acetals with 1 3 positions these 2 hydroxyl groups are in one true position one carbon two carbon. So those are one two position if you treat it with benzaldehyde, it prefers to form the acetals in one 3 positions, which means the 5 prime and the 3 prime will be forming the acetyl.

So this will remain open and this will be forming. So therefore, the 5 prime and 3-prime will be blocked and now you can do the chemistry on the hydroxyl group in the two prime positions for

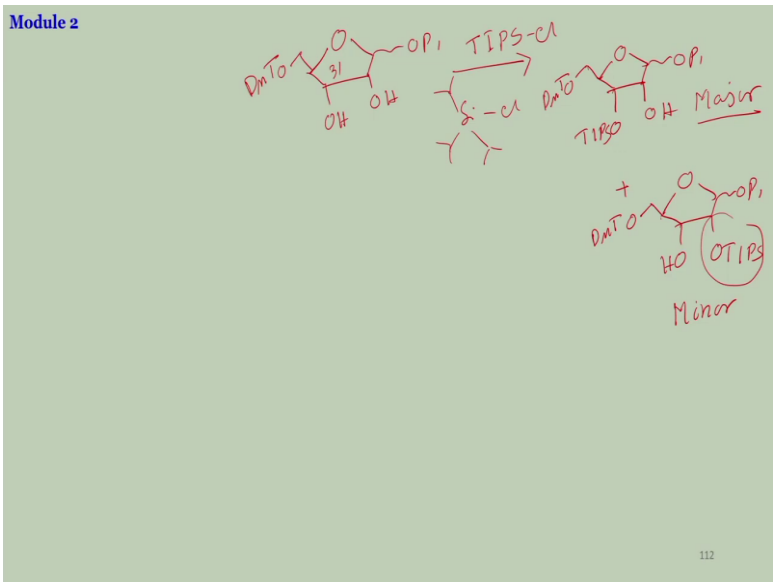
example if you want to do TMS chloride trimethylsilyl chloride then, you get will get OTS. Or you want to make acetelation with acetic anhydride various protection group protecting groups you can use.

So this will be OTMS or it can be OHC and then O this will still remain as protected and then you can open this up. So I am writing you away OP2 this will open up, this one open up so you can get a selective protection of, on the two prime position. So this is kind of an old technique what little bit advanced technique is that you use a Sallie compound specifically TI PDS chloride this is a TIPDS chloride.

Thus I will show you the structure. The structure is it is a crowded compound this is isopropyl, this is isopropyl O, this is Sallie, Di isopropyl and the chloro. So because of its large structures this also on reacts with a diol but again one to diol does not work here because they are too crowded. So here also the preference is 1/3 diol. They will react so similar to the benzaldehyde in this case.

So here you will get elimination of chloride and the C3, you will get the elimination of the chloride. This OH will remain free O, Si O, Si O isopropyl, isopropyl. This one will be forming this will be forming and your OH at the two prime position will remain free and after this you can do the same protecting group chemistry that you have done here so it will give you the Selective protection as you see this. And later on you can also open this up just by treating it with a slight amount of acid or any fluoride.

(Refer Slide Time: 07:50)



And now if you want to do selectively put selective protection of the C 3 prime this then what to do? So here there are some chemistry that have been developed. So one is that if you start with DMT on so which means already you have to do two steps one is here protection here second one is protection there. And if you treat these with TIPS chloride which is this Sallie, tri- isopropyl Sallie chloride, basically so this is also a crowded with high steric crowding Sallie compound.

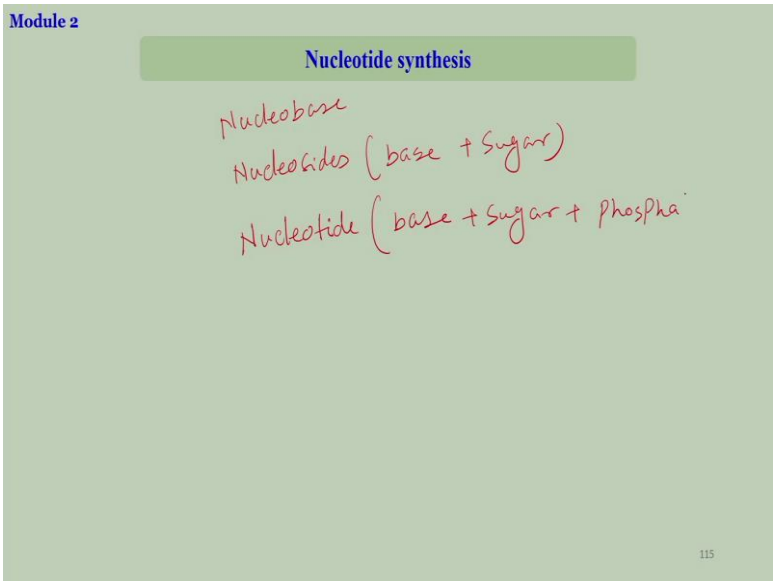
So if you treat this here obviously there will be nucleophilic substitution. Now question is which way is we will react this CH can also react, this CH can also react. You get both the compounds. One is if you treat it with a selective equivalent amount, then, this would be O TIPS. This is your DMT plus you get the two prime protection also.

PI yeah PS this can be free in this case so either this or either the 3 prime can react or the two prime can react now because of this bulky substitution. Here the reaction at the two prime positions is less much this is not the major product; this becomes the minor product because the steady crowding at the two prime positions is maximum. This is sandwiched between the 3 prime as well as the 1 prime very close together.

So the accessibility or the nucleophilic substitution by a bulky group is difficult in this position compared to this. So most often you get a major product for this but nevertheless you do get a mixture of these two compounds but this becomes a major product and that becomes a minor one. If you really want the two prime position to be selectively protected you have there are other ways to do it. You can figure them out. You can do individual protection, reproduction.

And likewise you can play around with the same protection and the protection chemistry we have already done here and then can find out a proper method but that will take that will require more number of steps. I guess so this is one of the ways people nowadays use for selective protections of the 3 prime hydroxyl group. So now this is actually for the nucleotides synthesis and we have already done and this is for the sugar chemistry. Now today I will come back now to the nucleotide synthesis.

(Refer Slide Time: 11:11)

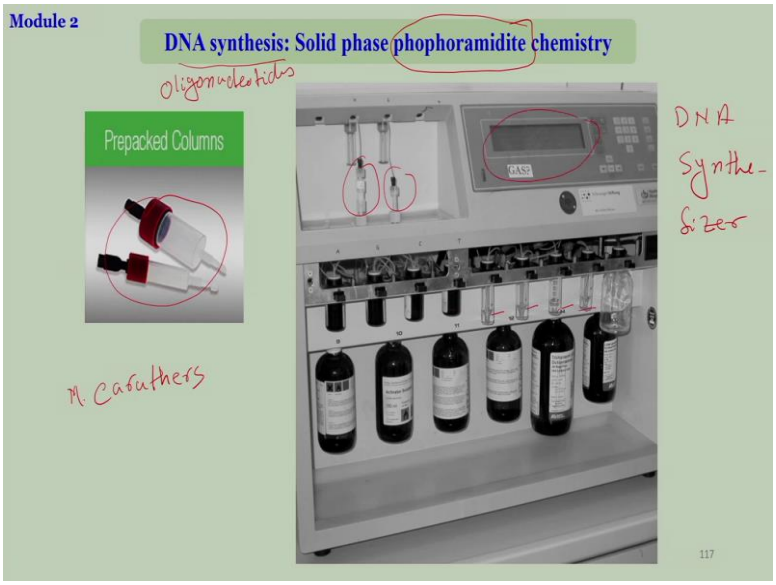


So we have done nucleobases then, we have done nucleoside that is basis plus sugar base plus sugar. That is nucleoside and now comes the Nucleotide base plus sugar plus phosphate. Now here, I will not show individual nucleotide synthesis. We will straightaway go for the synthesis of a chain of DNA. So basically nucleotide synthesis includes this is one of the key aspect of the chemical biology and enormous contribution from organic chemistry that has enabled us to synthesize short stretches of DNA in the laboratory.

So, small fragments of DNA cannot be obtained in the biological systems. In biological systems all we have a long genes and the chromosomes to work in your laboratory to understand the role of specific sequences in for biological functions as well as for there are plenty of other applications that we will see just in this next module, if you have a fragmented DNA. So synthesizing short stretches of DNA typically it can be any length may be 5, 5 base pairs to up to 50.

We can go even longer now it is base pairs DNA they are very important for the working in the biological laboratory or chemistry laboratory. So we will see now how this DNA can be synthesized in the laboratory.

(Refer Slide Time: 13:35)



And this is a beautiful chemistry that has been developed by M. Caruthers, he is an American it is not the same Caruthers who wrote the organic chemistry book. This is Marvin Caruthers who is famous for the synthesis for developing complete protocol for the synthesis of DNA stretches which is known as Oligonucleotides. So this is basically when you talk about DNA synthesis we basically means Oligo nucleotides which means the oligomers of the nucleotides.

So that involves the synthesis of the nucleotides as well. So the chemistry is a little bit modern organic chemistry and in advanced level. So I will show you the individual steps just in a moment. But still for this course I am teaching this because in this synthesis lot of different aspects have been developed. And they are really interesting one fact is that that this whole of this chemistry is done inside a machine.

So you do not have to go to a laboratory to your laboratory and mix things up in a into a round bottom flask or in a test tube and do the chemistry. So, everything is can be performed within a machine and that machine is programmed in a, with a computer. So just by putting the software on in place you can comment and then the machine will do the, do every artwork for you so this whole chemistry involves 6 seven different steps that I will show now.

But all of those steps happen one by one step by step commanded by the computer in this machine. Here you can see the machine. So this chemistry is known as the phosphoramidite chemistry. So as I mentioned the organic mechanisms or they are actually a little bit in advanced level. But still I would like to teach it even for the school children, just to show that how chemistry can be done, inside a machine inside an automated machine that can perform by

itself. So this is one of the machine that I am depicting here the machine is known as DNA synthesizer

This is a typical model of DNA synthesizer which had bottles here you can see there are many bottles here that are attached there. And these are the bottles for reagents, the different reagents that are needed in every step. And there are also these are little bit larger bottles because you need them in, in millilitre quantity for the synthesis and then there are smaller ones this-this-this you can see they are very small actually for which you need the volume in the level of micro molar quantity.

So we can do very fine chemistry here in the level of micro litre volume of material can be or liquid can be used for the synthesis. These are the other reagent bottles which are little bit larger which require a little bit more quantity than this but much less than this. And these all these bottles are connected to two channels here you can see the channels and those channels go all the way off here here as well as there.

And this is where your synthesis will, happen actually. So this is the place where the reaction happens and these are the vessels they are called columns or called the cartridge basically they are the, I will show you what are these, these are basically cartridge cartridge which has some solid materials inside. So these are placed in this chamber and the Machine pumps of all the required reagents in adequate volume all the way to here and the chemistry happens here.

And there is a display over here which can tell you what is happening, which step you are in or some monitoring steps are also there. So this is a beautiful invention beautiful discovery that allows you to do multi steps organic reactions, very in a very clean manner all automated. You can you can simply press the start button in the computer and go home. And the machine will do all the work for you.

(Refer Slide Time: 18:22)

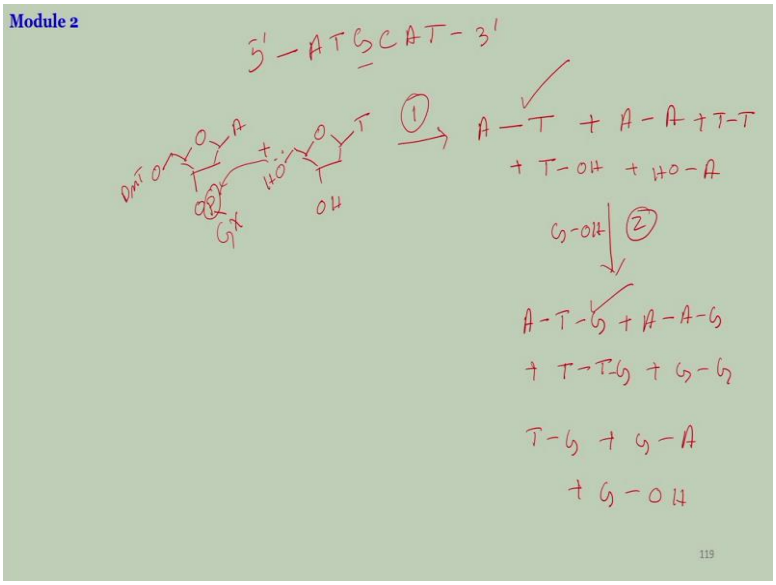
So you can all you have to do to tell the Machine that see that the desired sequence or your DNA that you want to synthesize and then it will do it automatically. So there are too many machines here. These machines are very much commercially used by the pharmaceutical industries as well as the biochemistry industries which supplies and the oligonucleotides to customers.

These basically synthesize a lot of DNA for different purposes which we need for our laboratories so we can purchase from them. Many number of machines connected to computer as you can see, so all automated. So now, the synthesis here, the synthesis protocol involves a solid phase synthesis that is what I have written here DNA synthesis is done in a solid phase force for immediate chemistry.

Phosphoramidite is the name of the reagent that you use for the nucleus or it is kind of a derivative of the nucleus. You do not have to remember or you do not have to even understand much of the mechanism that I will show. But still it is I think it is interesting to see how a stepwise chemistry can be done and how clean are they. And the whole chemistry is done not in a solution phase but in a solid phase.

So the solid powder is is packed inside this cartridge. So if you do, solution phase chemistry, why not this is happening in solution? So, if you try to do solution phase chemistry in a in a round bottom flask in your laboratory, or in a test tube, so for example, let us take a target first.

(Refer Slide Time: 20:02)



For example, I want to synthesize a random a DNA sequence A T G C. It is ATGC and AT, 5 prime, 3 prime. This is my target that is what I want to synthesize. Now, if you want to synthesize in solution what you will do first you will take A. So this is your A basically there will be hydroxyl group here. And here you will see you will require a DMT that you will see a protection would be needed because you do not want to touch here 5 prime.

This is your A+ you have to take your T maybe this is OH or it can be other protection group whatever chemistry you do. I will show you the problem here. So and of course you have to synthesize the phosphate so what you need a phosphate and then an X whatever that can be it can be a living group so because you need a phosphate. I am writing p4 phosphate and X for a living group.

So the essential chemistry is this will go and X will be eliminated. Let us say this chemistry works although it is very hard to find do a direct chemistry like this. But let us say this works. So if you do this chemistry then what you get. You get A T sequence I am not right drawing the phosphate so a phosphate T basically in solution phase you get your desired A T so first is a tea you get your A T. What else do you get?

One A may react with other A that will give you A, A. One T may react with the other T that will give you T, T plus. You will have unreacted T I am writing hydroxyl group free because of this plus unreacted A, not all reactions are complete. So if even if you use when you use these reagents after the end of the reactions they cannot complete all of them will not be consumed.

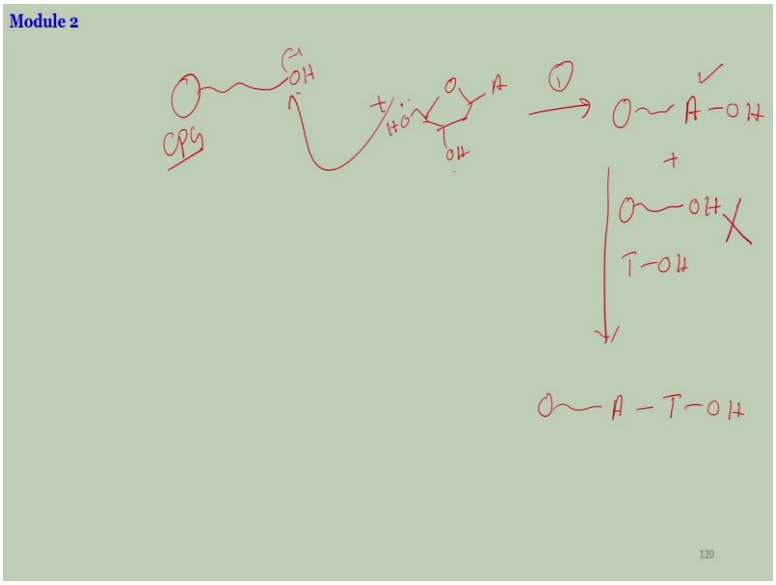
So chances are high that some of these materials will, will remain as in free form even after the reactions. So after the first very first step what you get you get 1 2 3 4 5 components in your solution. Out of this, only this one you desire. So if you want to get this in pure form you have to separate all others. That is again that is a difficulty or you can move on to the second step second cycle.

And then use G. G OH means I am writing because since the 5 prime is reacting always I am writing OH for 5 prime position, GOH. Of course, the G means whole of G. Here only the replacement will be. In the second phase what will happen? You will have A T G your desired compound plus here A A G plus T T G Plus G may react with itself you will have G, G you had some unreacted T left over here.

That can react now with this G in the second step. So we will have T G you had unreacted A which may again react with G. So we will have G A and then again unreacted G will be there. So it after the second steps you have multiple, so many number of products that are formed out of them. Only this is your desired compound. So if you carry on, if you, like this, if you want to synthesize just 6 membered ring a 6 membered DNA then, the number of products would DNA products would be so high that you cannot find even your desired component.

Or else you will need a purification in each step which is really a painful job. So when you want to synthesize 50 more long DNA 50 equally bases long DNA then obviously you can see the problem in the solution phase. It is very, very hard to do in the solution phase rather if you choose to do it in a solid phase. Solid phase is like this.

(Refer Slide Time: 24:58)



Solid phase synthesis means you have a solid support. In this case we use glass support. So this is a glass basically CPG is called the glass basically glass powder and that is connected with a functional group. Let us say it is connected with Ted O just consider kind of a DMT. Now this is solid so you treat now this with the first nucleobase was eh-eh-OH-O. So, instead of I am straight ahead taking OH here because I have to deprotect DMT otherwise again, it would record an additional step.

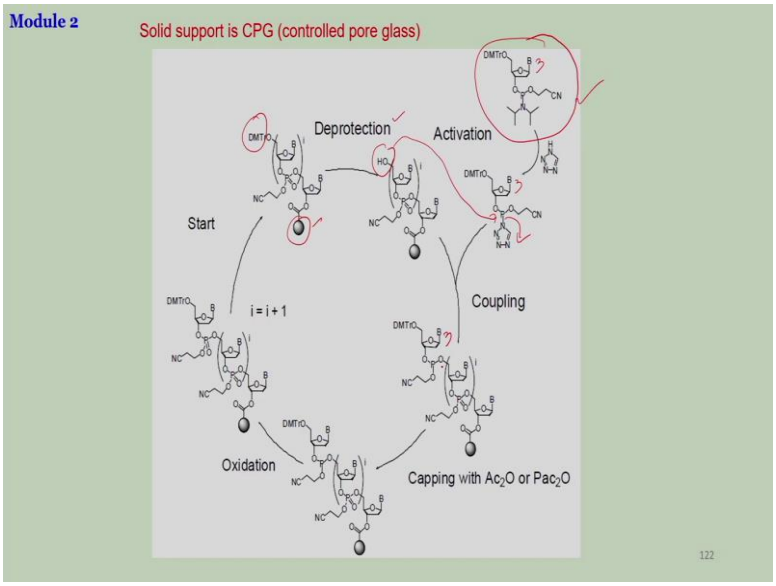
So this is connected this functional group is connected to the glass solid support now with a proper reagent. You can do this reaction and you can have A here with the OH free at the 3 prime end. And what else you can have? You can have the unreacted of this. But this is now in liquid form. So after doing the reaction if you wash away all the liquid and only get the solid then all you have here is this plus of course the glass which did not react, the unreacted glass.

And now I will show when we actually do the cycle of the DNA that we actually make this one inactive so that it does not react in the further step. Now in the second cycle, so this is the first cycle, second cycle is you add the T, do the same thing, you get A T OH free and of course, all the unreacted liquid form can be washed away. So if you go like this and obtain your solid all the time, the chances of getting the side products or undesired products, smaller ones, are very, very less.

So the productivity of the synthesis would be much higher if you do the solid phase synthesis because you are doing it in a located place. And you can remove the solid thing out of the reagents which is in the liquid form. So that is what the beauty of this synthesis, that this

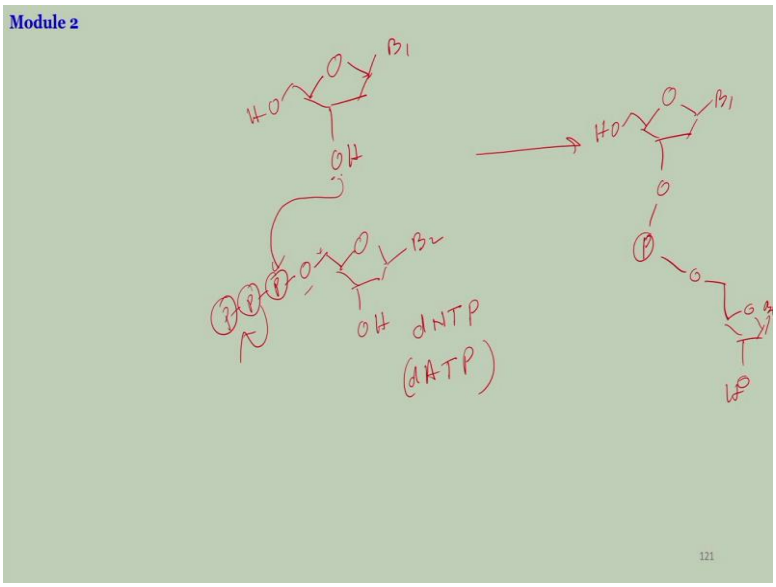
particular the phosphor emanate chemistry works on, solid support and it is a solid phase DNA chemistry.

(Refer Slide Time: 27:37)



So now if you look at the biology how these nucleotides are synthesized, how this phosphate is synthesized, how that 5 prime position and the 3 prime position of a nucleobase react? B1 B2 OH.

(Refer Slide Time: 27:47)



In biological cell, it is a spontaneous process and this reaction happens in presence of enzyme that we will see in vivid details in the just in the next module. So how it happens is that it has the Triphosphate O P P P I am not writing the phosphate. This is Triphosphate which is called the in general they are known as DNTP. Deoxyribose in its nucleobase T is Triphosphate for example ATP deoxyribose ATP is a form of triphosphate if there is adenine here is called

adenosine triphosphate. So, in our body we have triphosphates we have a nucleobase triphosphates, they are available.

And this reaction is catalyzed by the enzyme polymerase we will see just in the next module reacts with the phosphate eliminates diphosphate. And that will give you the nucleic acids phosphate so, B2 B1 OH that is how you get it in the in the biological cells. Now you try to do this reaction in laboratory you will not succeed. It is very hard to do this reaction without the use of the enzyme. And therefore, we had to develop, people had to develop for a very long time there are many different methods that make changes to these variations to make this reaction happen.

So in the solid phase for phosphoramidite chemistry we will see that what is the reagent that we use instead of that triphosphate it is this nucleobase analog that is used. This is the base here they are 5 prime is protected with a DMT and the 3 prime is protected with this which has a phosphorus group. And this is because this chemistry works well. It is only for the laboratory and this, after a lot of research and a lot of modulations, people have come with the idea that this molecule works best, so far.

So here I will show you the first the cycle how it works and then go to the individual steps when you do it in the machine in the DNA synthesizer. So, the first thing that let us start with here here is your solid support CPG and you have already done a little bit of synthesis. And now we are trying to add more nucleobases here. So the first step is that you deprotect the DMT out of this. And convert it this DMT is converted into the free hydroxyl group here.

That is the step called Deprotection. So this is attached to the solid glass and then the reagent the second nucleus that you wish to add will pour in, will be pumped into the solid. And then, it will react this OH is free and it can act as a nucleophile so this who can react with this phosphate. Before that this molecule is stable so you need to activate it. And that is done with activating reagent like this.

So it will activate and then this OH will react to this phosphorus clipping this out. This is the coupling means you are joining to the solid phase with the new nucleobase. This is coupled together. So if you turn this as B3 let us say, here is your B3, so this should be B3. Your B3 is

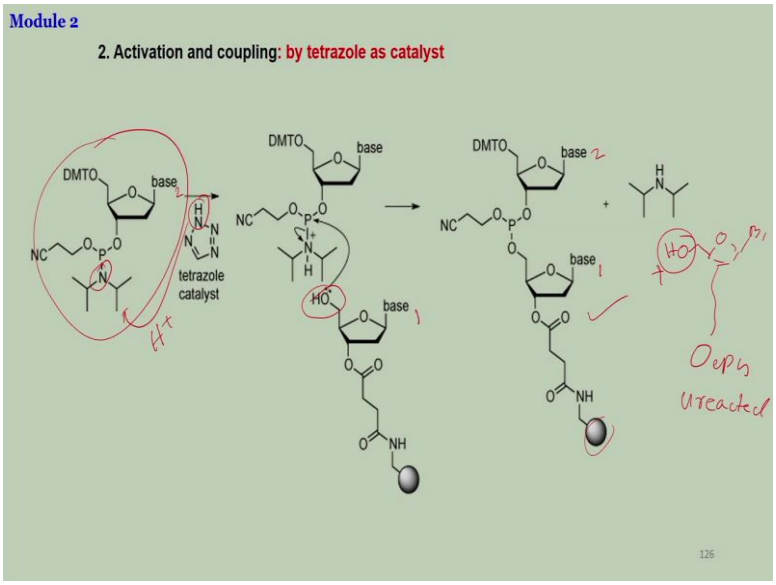
now attached with the solid phase along with the others. So coupling is done, your essential part is done here.

And now, if you see as I have talked about that, you will have the unreacted material here the unreacted solid which did not couple will be left over. And if you leave it as it is then it will react in the next phase. We do not want that. So therefore, the unreacted thing is capped and gone away once that is that unreacted part is capped, it is dead. It cannot react again. So after that what is left you get the free form the same thing without the impurity without the other solid impurity.

Next, if you see here, you are in your phosphate the phosphorus is pentavalent and that has a oxidation charge of plus 5 on the phosphorous. Here if you see that we have started with the plus 3 oxidation state on the phosphorous. Here also you have a plus 3 oxidation state of phosphorous. So we need to oxidize this phosphorous further into the plus 5 state that is called the oxidation. What it does is that it adds a double bond oxygen here. So you now have a pentavalent phosphorous.

And then you can continue with the cycle. You can still see your phosphorous is not really ready you still have a short tail and that we actually can eliminate in the last stage that I will show.

(Refer Slide Time: 33:38)



So I will take the individual step one by one and then you can see how it over how it happens. So first let us start with the first nucleobase attached with the glass. This is your CPG solid, it is

basically, solid glass that is connected via a linker. And that linker is connected with the forced sugar. This is the first base, base 1. And 5 prime position is protected with DMT. So, the first step is that you have to eliminate the DMT.

And that is done by using a very weak acid that is trifluoroacetic acid or trichloroacetic acid just a little bit of 3% very dilute solution of this, it will eliminate the DMT, will make a free hydroxyl group. Again you do not have to remember all the mechanisms, but I am just trying to the mechanism is shown here but I will not explain it in great detail because I want to show you the exact processes first. That is the essence of it.

Usually I teach this in an advanced level so that is why I want to explain the mechanism in detail. So you get a free hydroxyl group here and your DMT will be eliminated in the form of a carbo-cation and this has a beautiful colour. This has an orange to purple colour and that is actually a way I will show you here, to understand whether your reaction is happening or not. Here once it happens here, you can see the outlet chamber when the liquid is going out of it, you can see the beautiful orange colour that is passing through the channel.

So that proves that your DMT has been deprotected; it is kind of monitoring the reaction actually. DMT will be out and you will have free hydroxyl group. This is the most crucial step actually and now you have this free hydroxyl group with the solid phase attached base one. Now you add you pump in your phosphoramidite this is your phosphoramidite this is base two with the DMT protection of the 5 prime end.

And this molecule is pretty stable and therefore is less reactive. So we have to make it a little bit reactive so that this hydroxyl group reacts there acts as a good nucleophile. What we do is we use a little bit of Tetragyl. This acts as a catalyst here what the Tetrogyl does is that Tetrogyl is just a little bit more acidic compared to this nitrogen. So this nitrogen group here is little bit acidic compared to this nitrogen. And therefore it provides a proton here.

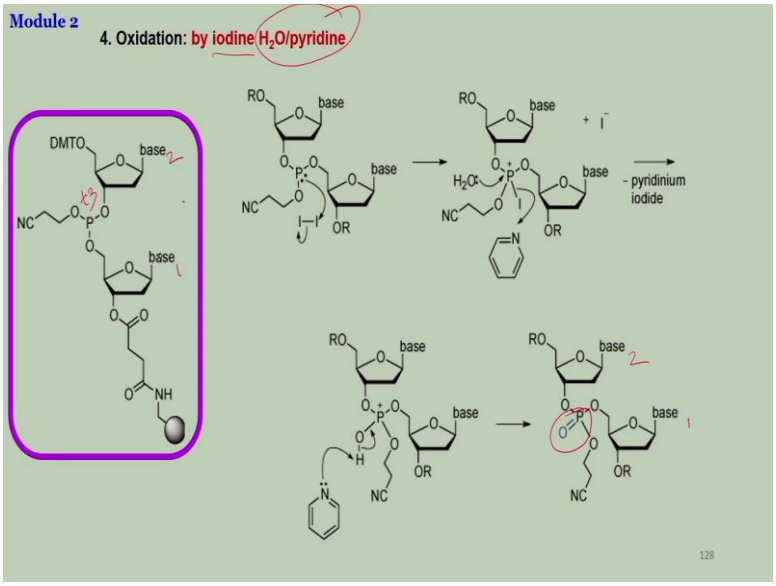
The proton from the tetragyl is shifted to this nitrogen. And then, you have the NH plus in this position. And Tetragyl becomes n minus which can be stabilized by the internal resonance. So essentially you have this form and this becomes a good leaving group. So now the hydroxyl group can react here eliminating this. And you have this compound here not the compound with

the solid phase with your base one and this is your base two plus you have the solid with the base one.

I should draw this is base one and your free hydroxyl here. This, this, this, this, this whole thing which did not react, unreacted this, this is your CPG, unreacted this. Now you have to make this inert otherwise it will react in the next phase. So first thing what we do is we wash away all the liquid out and write. So all these unreacted of the starting material this is gone; all the by product that is in liquid form are all gone.

Tetragyl are gone, all in liquid phase are gone. What will remain is only this and this, the unreacted one. Now we need to cap this. We need to make this inactive. So that is called the capping stage.

(Refer Slide Time: 37:44)

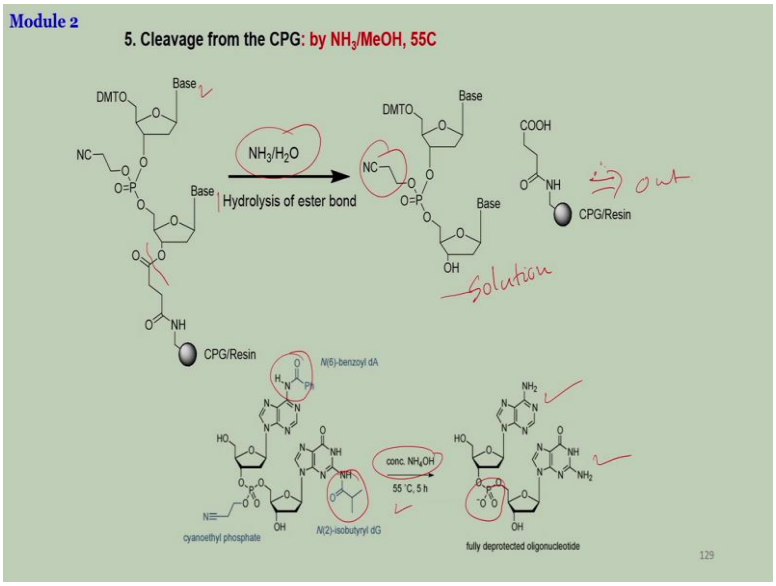


So this is will be capped is basically a stratification reaction you will form an Ester here. Once you form the Ester is very unreactive it will be dead. It will not react in future steps. So all you have now is this with your base 1 and base 2. Now if we want to stop here also what are the other steps that are remaining? If you want to stop right here, if you want to continue, you can move on just like the two steps that I have just described.

So now here as I was saying that it has a +3 oxidation state so it needs oxidation which is done usually by iodine in presence of water and pyridine. All it does it includes a double bonded oxygen in this position. This is base 1, this is base 2, OR and this OR means the protection

groups. So you have basically 5 plus 5 oxidation state phosphate ready. So we, you have still have this tail.

(Refer Slide Time: 38:48)



So now, then, once you complete all your synthesis what is your aim is to get rid of the solid phase because we want your DNA, not the solid. So that is done by hydrolysis with in ammonia water if you just hydrolyze in ammonia water then this ester bond will cleave and that will give you free from the resin. You will get a free hydroxyl group and your resin will be eliminated out. This is now in solution phase.

This can now come in solution phase solution and your solid is out you can take your solution and then can isolate from there. You can dry the ammonia out and dry the water out you have your compound and this group as well as this group can be eliminated to get the pure phosphate, if you just treat this with ammonium hydroxide at 55 degree Celsius for few hours, that will give you this.

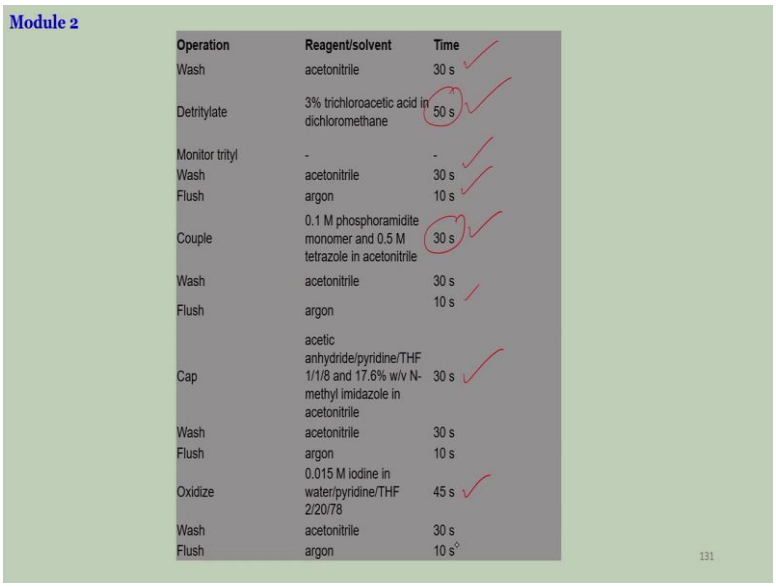
At the same time so one thing I forgot to mention that when you use the nucleobases for example the guanine or adenine those have free amine groups which may react, so that needs the protections. I have talked about the protections of nucleobases ah already, so you have to use at the beginning the protecting groups of the nucleobases base 1 or base 2 if they have free amines.

And at the same experimental condition those protection groups will also be de protected and you will get the free form of the nucleobases. So that iss how you get a dinucleotide and you

can carry on from this stage itself you can carry on if you want multiple numbers. And that is how we can go really far. At this point now 30-40 nucleobase pairs are very easy to synthesize. We can go be even beyond 50. So that is how nowadays the for the DNA is synthesized.

And they are called the all uniquely tides and they are used in for various various purpose you will see there are huge applications just in the next module. So this is the again the cycle you can go around the cycle again and again and it will give you this.

(Refer Slide Time: 41:05)



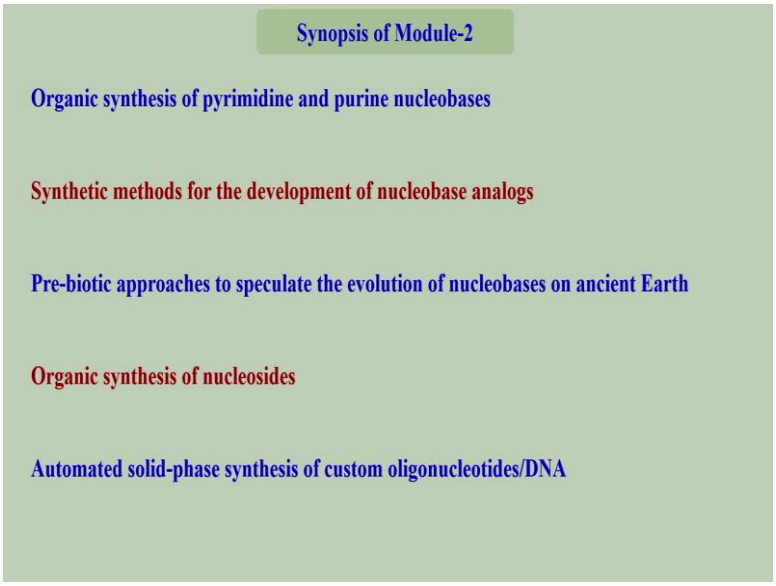
So here is the time frame how much time you need do you need to complete this. You say it is really fast, really beautiful chemistry, a really fascinating chemistry and every step we have washed to remove all the liquids out. So, first step is washing with the acetonitrile that takes only just 30 seconds. Detritylation that is the removal the DMT just take 50 seconds then wash with acetonitrile again that takes another 30 seconds.

Then you make your solid phase dry because so that no other liquid stays in. That takes another 10 seconds with a pure organ and then coupling where you attach the new nucleobase with this that takes only 30 seconds maximum one minute 30 seconds to 1 minute power base coupling. And then wash, Argon flush capping takes another 30 seconds. All the washes that takes 30 seconds oxidation needs a little bit longer time 45 seconds and then 30 seconds wash and flush.

So within 5 minutes your reaction is done, full one nucleus one nucleotide synthesis is done in 5 minutes. So in one hour you can synthesize are quite a good length of your DNA and all automated inside the machine. So ye so that is what is called the phosphoramidite DNA

synthesis, automated solid phase nucleotide synthesis. And with this, I will complete the module two.

(Refer Slide Time: 42:44)



So what we have learnt in this module is that I have shown the organic synthesis of pyrimidine and purine nucleobases. We have seen synthetic methods different methods that we can use or develop for the nucleus analogs, if you want to synthesize the derivatives of the nucleobases which have other applications apart from the base pairing, for example, in drugs industry and medicinal in medicine industry as pharmaceutical drugs.

Then we have also seen how prebiotic chemistry can be done or can be some ways can be speculated, how the evolution of the nucleus has happened when there was no life on earth. Then we have seen the organic synthesis of the nuclear sites how can you couple the nucleus with the sugar both in RNA or in DNA. And then finally we have seen how full DNA or not a full-length it is we call it full-length DNA because that is how we understand.

It, it is Oligonnucleotide, how can we synthesize oligonnucleotide using our automated solid-phase synthesis. So that is all for the module two and the next lecture I will start the module 3. Thank you.