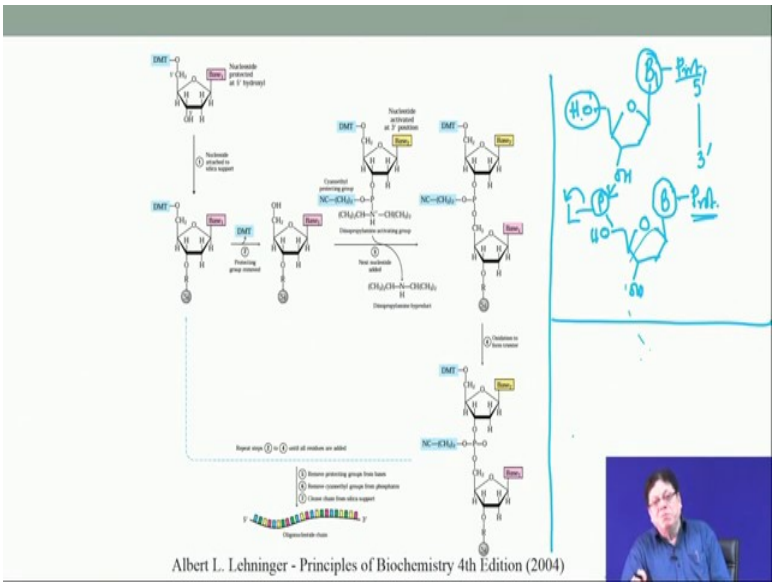
Organic Chemistry In Biology And Drug Development
Prof. Amit Basak
Department of Chemistry
Indian Institute of Technology, Kharagpur

Lecture - 27
Synthesis of Oligonucleotide

Welcome back to this ongoing course on Organic Chemistry in Biology and Drug Development.

(Refer Slide Time: 00:30)



In the last session, we have discussed the various methods of DNA sequencing and also covered one method called next generation sequencing techniques. Specifically we mentioned the chemistry involved in pyrosequencing.

Today, we will discuss the chemical synthesis of Oligonucleotides; that means how DNAs can be synthesized chemically. In case of synthesis of peptides, we have to consider several aspects because the amino acids have free NH_2, free carboxy and occasionally reactive side chain. So, we need protection deprotection strategy. Protecting groups are placed on the amine of one component and the acid of the other component. The remaining amine and the acid functionality of the 2 amino acids are coupled. So, that is called a coupling stage.

So, basically there are three stages- protection, coupling and then deprotection. You cycle this and then finally, through a common deprotection step you strip of all the protecting groups from the molecule. Initially all the synthesis were done in the solution phase. But later on people realized that the length of the peptide cannot be made in a very large in solution phase. So, Robert Bruce Merrifield developed the solid phase synthetic methods of peptides where the amino acids are linked to a solid phase and all the reactions were done on the amino acids that is anchored to the solid phase.

Now let us discuss the synthesis of oligonucleotides. I will first write the structure of a mono-nucleotide and then try to analyze that what are the possible problems and how to solve those problems. We are talking about deoxy oligonucleotide for the synthesis of the DNA and this is the general structure of a nucleoside.

We are planning to couple this with another nucleoside. But the connection between the 3 prime and the 5 prime OH is basically a phosphate linkage. Here coupling means the coupling of these mononucleosides. But there are number of OH functionalities very similar to the peptide case. So, what we have to do? We need a protection of this if we want a reaction between this and that.

It is connected to a reactive phosphorus moiety and this OH then attacks the phosphorus. There must be some leaving group attached to the phosphorus and that leaves. So, the principle is that you should have protection of one of this OH. Remember, that DNA synthesis takes place enzymatically in the 5 prime to 3 prime direction. But, in chemical synthesis you have two choices either your 3 prime OH can attack a 5 prime active phosphate group attached to the 5 prime OH or in the other way, 5 prime OH attacks the reactive phosphorus moiety attached to the 3 prime OH.

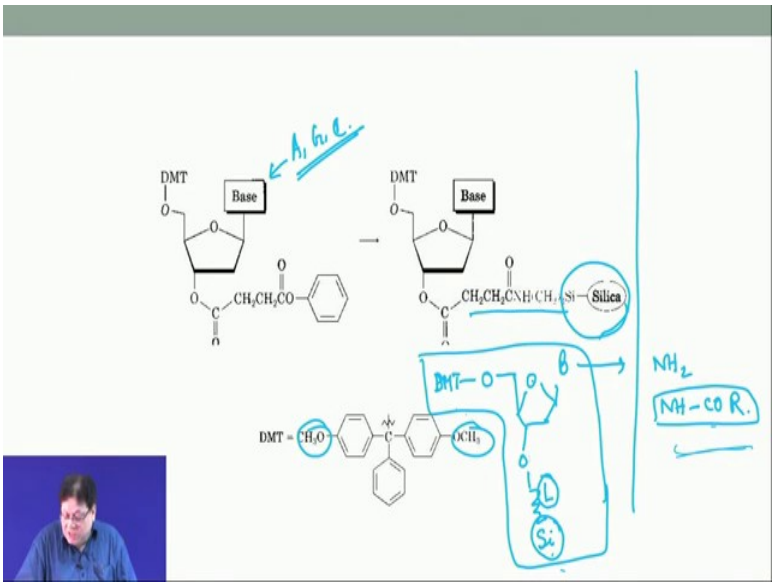
So, now the question is by which way the synthesis is generally done? The synthesis is generally done by the solid phase synthesis method because of the advantage of purification and also large number of nucleotides can be added one after another.

If strong nucleophilic groups are present in the bases it will hamper this nucleophilic attack on reactive phosphates. The 3 bases *i.e.* Adenine (A), Guanine (G) and Cytosine (C) contain free amino groups.

These amines are stronger nucleophiles than OH. So, you have to protect these amines of bases except thymine. Thymine does not have any NH_2 and that does not need any protection.

This is little bit more complicated than the synthesis of peptides.

(Refer Slide Time: 07:32)

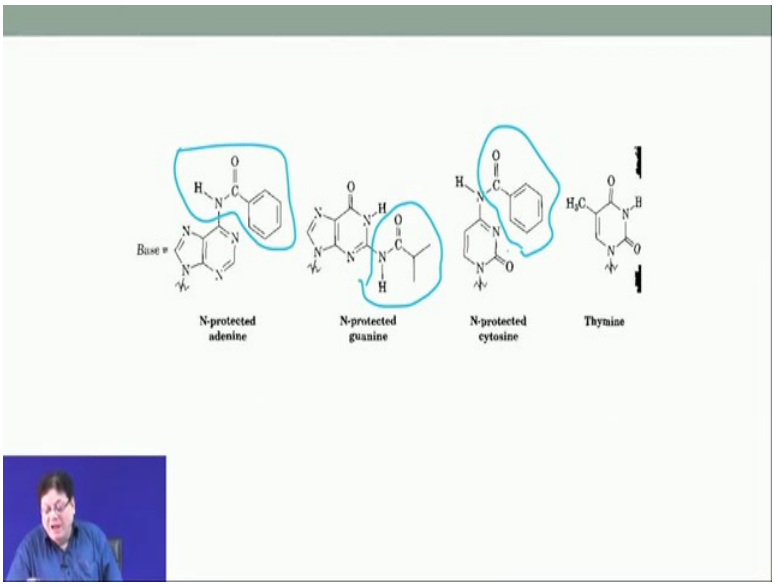


In the first mononucleoside, there is no phosphate. If the base is A or G or C it needs to be protected. I will show what the protecting groups that are used here are and then this 3 prime OH is attached via a linker to a solid silica surface. So, this is your solid surface.

This 5 prime OH is protected as a DMT group *i.e.* dimethoxytrityl group. In this case, there are 2 OCH_3 groups. So, that is called dimethoxytrityl and this is hooked to a solid surface.

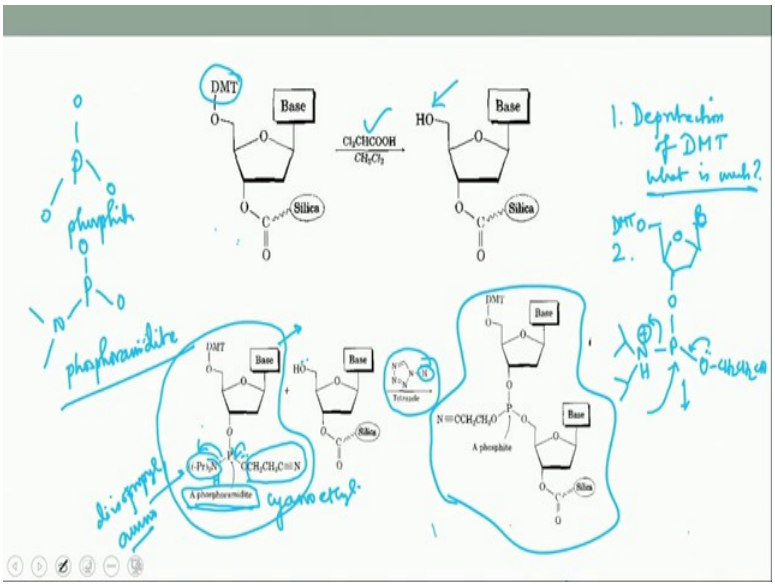
The general protected groups of bases are nothing but NH-COR. So, you do either benzylation or you do some acetylation. This 3 prime OH is the starting point and it is attached to the silica surface.

(Refer Slide Time: 10:14)



The next thing is the protection of base. Adenine is protected as benzoyl, guanine is protected as isobutyryl and cytosine is protected as benzoyl. Thymine does not need any protection.

(Refer Slide Time: 10:40)



This 3 prime OH is attached to silica base and this is 5 prime OH is protected as DMT. DMT group is deprotected by trichloroacetic acid or dichloroacetic acid. Trichloroacetic acid or dichloroacetic acid is little bit stronger than acetic acid. You know that

attachment of halogens to the alpha carbon of acetic acid increases the strength of the acid.

You can have another nucleotide which you want to join to this to this 5 prime OH. This nucleotide has base and as I said if it is A, G, C that will be protected as the acetyl or benzoyl groups and 5 prime OH is protected as DMT. The 3 prime OH is now is attached to a phosphorus and this phosphorus is attached to a tertiary amine group and this is also attached to a the other group is called cyanoethyl group.

This is diisopropyl amino group. If phosphorus is attached to 3 oxygens that is called a phosphate. If one of the oxygen is replaced by nitrogen it is called phosphoramide. So, that is known as phosphoramide, but the nomenclature is written as phosphoramidite in the books.

This is the free OH and this one is attached to protected 5 prime. 3 prime is attached to a phosphorus which has a very good leaving room because the diisopropylamine in presence of a proton source will be protonated and as it is protonated it wants to leave because the oxygen lone pairs are there. So, the oxygen lone pairs can fly and kicks out the protonated form of this isopropyl amine group; that means, you are kicking out isopropyl amine ok. Here tetrazole is used as slightly acidic group.

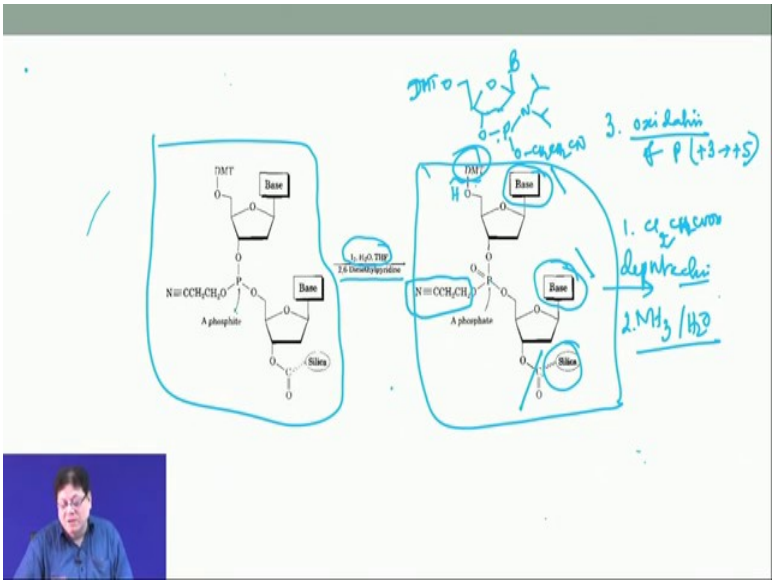
This hydrogen is quite acidic in tetrazole and it is believed that this hydrogen protonates the nitrogen of this. So, this becomes NH^+. Then this oxygen goes here and this diisopropylamine goes out. This oxygen becomes plus and then this free OH attacks this phosphorus and forms this is phosphite bond. Now you do not have that plus 5 state of phosphorus, I said this is phosphide. You have phosphorus oxygen and this is a lower oxidation state of the phosphorus.

The mechanism is like that you have O-P-N isopropyl and I said that this becomes positive and this is O-cyano ethyl. Once this is protonated, then possibly oxygen lone pair will help the expulsion of the free amine. This is a base and remember this 5 prime OH is protected as DMT.

You have this free OH on the other component. It is your starting point which is attached to the silica and now that this phosphorus the 5 prime OH resulting in the formation of this phosphite. I hope this is clear. This is your coupling stage. This is called

phosphoramidite method and this strategy was developed by an Indian born scientist named Professor Har Gobind Khorana. He was the first person to develop the synthetic protocol of this nucleo oligonucleotide.

(Refer Slide Time: 17:25)



So, after the first coupling you have base protection except for thymine and you have a DMT protected here. This is a phosphite and this the 3 prime OH attached to the silica. If you want to extend it to a trinucleotide you can extend the step. The next step will be basically you have to first oxidize it. This is not the correct oxidation of state of the phosphorus. So, first you have to oxidize it and then oxidation is done with iodine THF water in presence of a base.

So, the third step is oxidation of the phosphorus from +3 to +5. It gives you the correct oxidation state of phosphorus, but still the groups are not this is protected. One of the oxygen is protected still and this is protected as DMT. Now, if you want to extend further then I have to add another one. After that I will take this DMT off generating the free OH and then I will add the DMT protected phosphoramidite group attached to the 3 prime OH.

If you deprotect this that will become OH and that will attack another nucleoside which is again protected as DMT. Here it will be O phosphoramidite. So, this is O-P and this is N of isopropyl amine. Here it will be cyano ethyl (CH_2 CH_2CN). Then tetrazole will

protonate this and then this OH is going to attack here and you get a trinucleotide in that sense. Let us see that how can we make the dinucleotide.

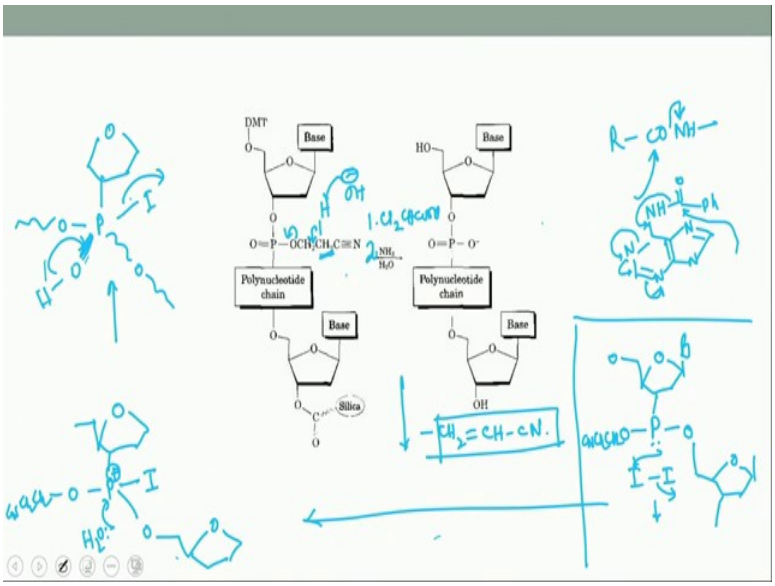
In the dinucleotide, you have all these protecting groups. The 3 prime OH is attach to silica that is a virtually a protecting group and then this base may be attached by benzoylation or acetylation the amine group. This is also protected as DMT and this is also protected as the cyanoethyl. So, if we increase it further we can do that, but the last step will be very similar to this.

Again I am repeating. The first step is basically you take a DMT protected nucleoside and then attach 3 prime OH to the silica surface solid surface. Then, you take the DMT off and add this phosphoramidite where it is protected as DMT. These are first two steps. In the third step, you oxidize this phosphorous to the correct oxidation state and if you want larger number of oligonucleotides you can repeat.

I want this dinucleotide. The next step will be just deprotection. You have to deprotect all these groups. DMT group is deprotected by trichloroacetic acid or dichloroacetic acid. These groups can be taken off by aqueous ammonia.

That base becomes OH, it takes up the benzoyl group at the bases. So, the bases are free and also that takes this group off. Again just a repetition, the third step is basically oxidation and the fourth step is called deprotection. If you want further extension you have to put the DMT off and then do the next coupling.

(Refer Slide Time: 22:36)



What is the mechanism of deprotection of base? In case of these bases, these amines are little bit special.

After leaving, the NH minus is stabilized by these 2 nitrogens. It is electron withdrawing nitrogens. These benzoyl groups can be taken up very easily because these are not as strong as the aliphatic amides. They can be deprotected very easily. This is not very difficult because these hydrogens of CH_2 group are quite acidic due to presence of the cyanide.

One of the H of CH_2H can be taken by OH minus and that goes in between this two carbon-carbon and that comes out. So, a phosphodiester is leaving and here acrylonitrile is coming. So, the byproduct is minus acrylonitrile which is a liquid and it can be removed since reaction is occurring on the solid phase. You always watch the bead, it will be detached from the solid surface at the last stage.

The mechanism that is left is the oxidation of phosphorus by iodine. There is a P-O and a cyano ethyl CH_2CH_2CN and you have here another oxygen which come from the 5 prime OH. The phosphorus is a nucleophile which attacks one of the iodine. One of the iodine is kicked out as iodine-iodine bond is weak.

This is phosphorus, you have this O-cyano ethyl and here you have attachment of another 5 prime group. In this case, the phosphorus is basically a plus. Then you have

water because I added iodine THF water in the system. Now, water will come and attacks this phosphorus because the phosphorus is positively charged.

This is phosphate; this is the cyano ethyl and this is the other chain. This is becoming OH as the water will just lose one hydrogen. Now it becomes the phosphate by expelling the other iodine. The whole thing is driven by the fact that iodine-iodine bond is weak. Other important issue is that phosphorous is strongly oxophilic. Phosphorus oxygen bond is very strong.

You know that in case of the reactions involving triphenylphosphine like your Wittig reaction or other aza-Wittig reactions, the end product is triphenylphosphine oxide. So, that oxide formation is very predominant in phosphorus chemistry. So, that becomes the phosphate. So, that basically ends up the solid phase synthesis of oligonucleotides.

Again, just have a quick brush up - the 3 prime OH anchored to a solid surface by a linker, the 5 prime OH was initially protected as DMT which is taken up by acid and then the you take another component that is the 3 prime OH is attached to a phosphoramidite and then the 5 prime OH is protected as protected as DMT. Now, in presence of a tetrazole the amine gets protonated and the 5 prime OH is attacking the phosphorus attached to the 3 prime OH.

You have to oxidize this phosphite into phosphate and then if you want further extension you take the DMT off from the top. Then you repeat the same set of reactions. Finally when you have the desired length, you take the DMT off with trichloroacetic acid or dichloroacetic acid and then you add ammonia aqueous-ammonia to deprotect from silica surface and protected bases. It also takes care of the cyanoethyl group which is deprotected as acrylonitrile resulting in the phosphodiester bond.

So, this is the current strategy to synthesize oligonucleotides of. Various biotech companies are actually supplying oligonucleotides of particular base sequence which is required. There is another molecular biology technique called polymerase chain reaction for sequence determination and you need primer. Primers are also synthesized and supplied by these companies. A lot of biotech companies now survive by just synthesizing oligonucleotides and sending it to different labs. One important point to note that our direction of attachment of one oligonucleotide to the other is from the 5 prime to 3 prime direction in case of enzymatic process. That means, the 3 prime OH is

attacking the 5 prime triphosphate, but here it is the reverse. The chain grows in the 5 prime direction.

So, this is the 5 prime OH attacking the 3 prime phosphoramidite; that means the phosphorus attached to the 3 prime. So, it is just the reversal of biological process. Similarly in case of peptide chemistry, carboxy side is attached to the solid phase. In solution phase you can do from the N to C direction or C to N direction. But, for the solid phase you start from the carboxy and make the peptides of desired sequence.

Similarly, here for the oligonucleotide you do the same i.e reversal of the biological process. That means, the 3 prime OH is anchored to the solid surface and 5 prime OHs are attacking the 3 prime phosphoramidite.

Thank you.