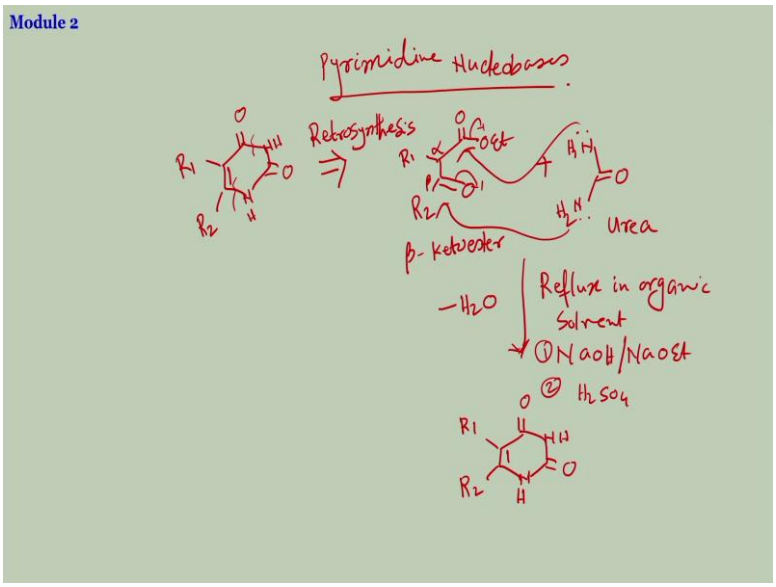
Essentials of Biomolecules
Prof. Kousik Deb
Department of Civil Engineering
Indian Institute of Technology dash Kharagpur

Lecture 07
Chemical Synthesis Purine Nucleobases, Prebiotic chemistry

Hello everybody and welcome back to the lecture. So, we are discussing the Synthesis of the Nucleobases and how the how you can manufacture the Nucleobases in the laboratory. And we have started with the synthesis of the pyrimidine nuclear bases.

(Refer Slide Time: 00:44)



Pyrimidine nuclear bases or there it is not only the pure nuclear bases also their derivatives because a large variety of those derivatives have shown interesting properties biological as well as the pharmaceutical properties. And we have seen in the last lecture that you can formulate a retro-synthetic procedure for the synthesis of such nuclear bases. And we have seen that if you have for example here a substitution R 1 can be any kind of substitutions are true can be any kind of substitution.

Mostly we will take the simple substitutions that we have talked about NH Co there is another NH CO here so this is a class of the Pyrimidine nucleobase uracil or thymine mostly and you can find out us Retrosynthetic procedure means going back. And we have seen that you can do it by cleaving these two bonds and starting with this kind of compound it can be either an ester or zinc chloride things like that with a good living roof at the place of O at R1, R2.

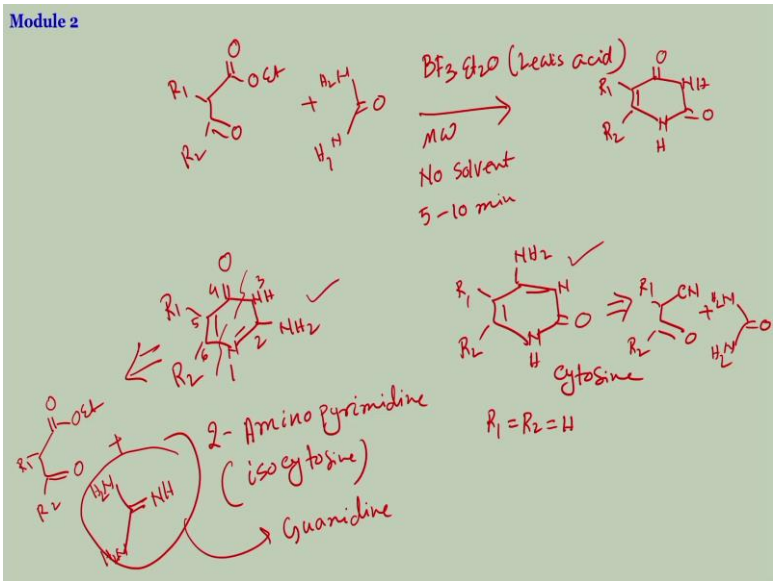
And this is a carbonyl compound plus very simple and readily available material that is urea. So in general this kind of compound is known as Beta - Ketoestar. So if you consider this is an ester this is the Alpha position this is the beta position so you have a carbonyl group at the beta position. That is why it is called the beta keto ester so a combination of beta Ketoestar and the urea. So this is the retrosynthesis.

Now, if you look towards the forward synthesis, if you start with this this Plus this and we have seen if you reflux in organic solvent suitable organic solvent actually that depends in that property of this compound in which organic solvent this will be soluble. You have to heat it up at reflux condition and in presence of strong base either sodium hydroxide or ethoxide followed by sometimes it takes takes strong acid also H2 SO4 then you will get the fusion product actually.

And you have seen that this is basically a nucleophilic substitution reaction. And this one was basically a ionisation reaction there will be OH and then followed by dehydration so it will lose one water molecule and give you the desired product. And this is what you are looking for this is what is your target, which is same as this. So the idea is first formulated retro-synthesis scheme and then with the starting points you have to manufacture, you have to come down to the target molecule.

So to verify or to validate whether the synthetic protocol that you have formulated is actually correct or not or is viable or not so that is why you have to do the forward process also and this is very much the process of the methodology or the technique that industries used or the pharmaceutical companies use to synthesize a variety of the pyrimidine compounds pyrimidine nucleobased derivatives in large scale.

(Refer Slide Time: 05:25)



Later on we have seen that you can further improve in the method or you can make the method more useful with more variations by a different, by a slightly different technique with the starting with the same starting material, same urea, but instead of the strong acid or the strong base and the use of organic solvents you could simply use you can replace all those by our Lewis acid BF3 which is present usually in ether.

So ideally it is the BF3 that is the active reagent which is essentially a Lewis acid which means the boron has empty orbital to accept electrons. And you can use microwave chemistry here and no solvent. You do not need a solvent here. Within 5 to ten minutes of reaction, you get your desired compound with really good yield actually R1 and R2 and you can have a lot of variations there.

So, the reason you do not need solvent is because of the fact that most of the times you are starting beta ketoestar, your starting material, this is liquid. This is in liquid form and urea is in solid form. So when you mix them together they are immiscible because urea is solid and urea is not soluble in this liquid. But once you heat it up once you start heating that up once you reach the melting temperature of the urea then it actually dissolves it gets melted and dissolves it into the liquid form of the beta Ketoestar.

So then they come in the single phase and therefore they can do the reactions without any organic solvent. So this is one case. Secondly, if you want to synthesize the little bit modifications which we have done as R1, R2 in here and Amino so this is called if you could

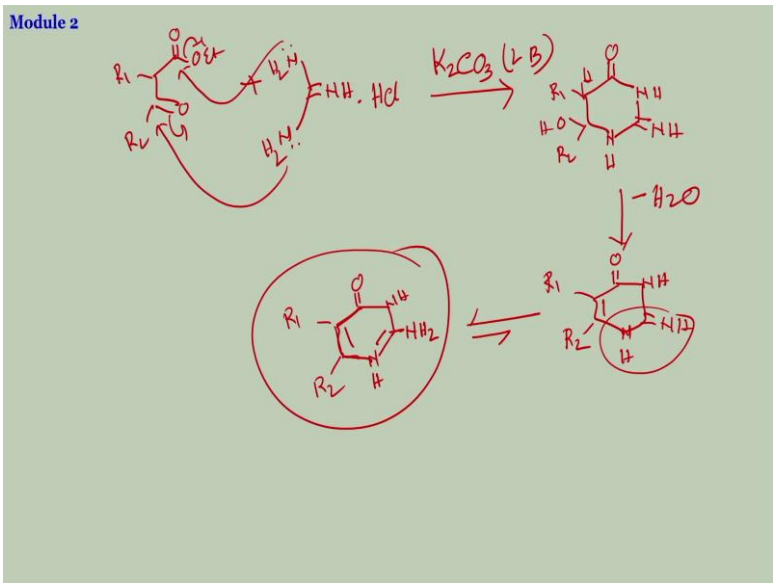
do the numbering 3,4,5,6 this is 2 Amino pyrimidine. And this is basically an isomer of cytosine if you see this is your cytosine here the amine is present.

So it is just that the amine group in the fourth position of cytosine has come down to second position and the keto group that the two position of the cytosine has gone to the fourth position. It is an exchange kind of thing so this is your cytosine so this is called ISO cytosine. So cytosine can be synthesized from the starting material here would be cyanide CN and this if you have R1 and R2 you will have here R1 R2 in general chemistry I am talking about for cytosine R1 equals to R2 equals to H.

Both are hydrogen this plus urea combination of these two under the similar condition as we have seen before will give you the cytosine. Now the variations and the unnatural one this is mostly unnatural one, Iso cytosine derivatives they are many of the medicines have this kind of structure, this kind of component in the total molecular structure of many medicines actually. So you can manufacture these by cleaving here and here.

It will come down to the same thing here beta ketoestar, beta ketoestar R2, R1 plus here if you see, this would be instead of urea, this should be, this molecule. Urea had one oxygen here in this case. You will have because you have to bring the NH2. So you will have the NH moiety here this is called this is known as guanidine. This is called Guanidine. So a fusion of guanidine with the beta ketoestar will give you the Two Amino Pyrimidine. That we have discussed already in the last lecture.

(Refer Slide Time: 11:09)



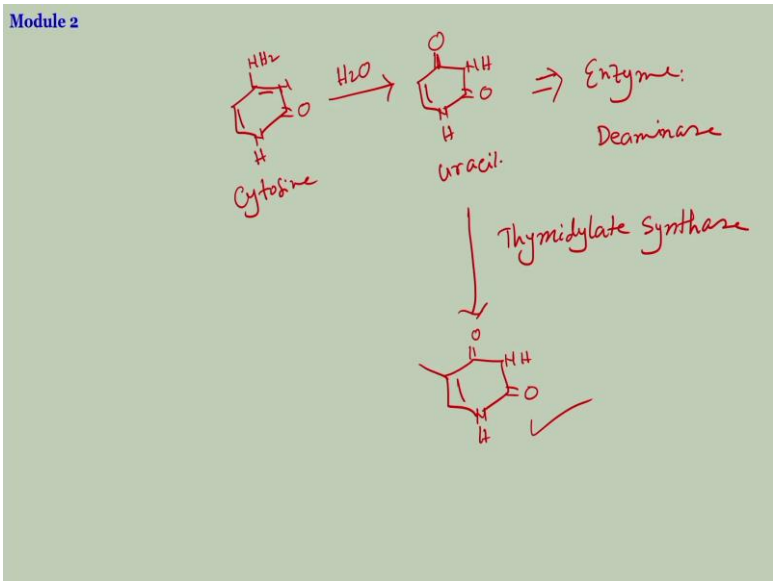
So if you want to do the forward reaction with this two amino Guanidines and this compound mainly exists when you buy commercially this does not exist alone because it exists in terms of salt most of the times in terms of the hydrochloride salt. So this compound exists as dot HCL. And when you fuse them together obviously first thing is that the HCL will otherwise interfere with your reaction so to remove the HCL.

Therefore you need to use a base in this but you do not need to use a strong base what we have figured out figured out. That if you use a mild ways just like bf3 was a Lewis acid potassium carbonate is a Lewis base L B, some mild base. So it can actually neutralize the HCL over time and then it will give you the same kind of an acyl substitution and this one is your hydration dehydration will occur here. So you will have this NH in here you will have the NH we have NH R1, R2 here will be OH here for there is H here so followed by minus water.

So I am showing with the forward process actually R1, R2. Now here you can see that this is the amine form so it will undergo tautomerism very quickly and will give you it's in h because this will be the driving force is the aromaticity yeah your target compound two amino pyrimidine or isocytosine. So likewise, you can make lot of variations you can change this into sulphur and here you can make sulphur so many variations that are present in our organic library actually actually and are commercially available also.

So you can you can find out the proper synthetic scheme depending upon your target molecule. So this is what you can do in your laboratory. By the way do you know how these pyrimidine nuclear bases are synthesized or biosynthesized in living cells?

(Refer Slide Time: 14:19)

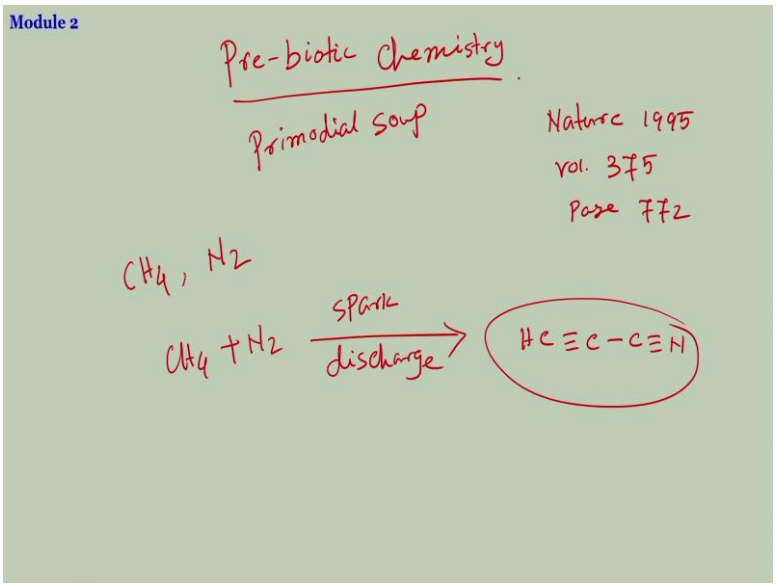


It is usually the cytosine which is first synthesized. NH this is your cytosine and then this is converted into uracil it is a pretty easy hydrolysis reaction basically. So this conversion happens with the help of an enzyme called deaminases. Here enzyme of course in biological cells everything is catalysed to enzyme. So in this case enzyme is deaminases, cytosine deaminases we call. So there are variations and so it is one kind of deaminase which will deaminate the cytosine removes the NH2 group and will hydrolyze to the keto.

So it is basically a hydrolysis reaction essentially an hydrolysis reaction that will convert the cytosine into uracil. So if you even if you keep cytosine in your laboratory in proper condition in across medium and little bit with a condition maintaining the conditions over time you will you can see this cytosine will be converted or will be hydrolyzed into uracil. And once uracil is synthesized this is then converted into thymine.

And this reaction I have mentioned before also this reaction is catalyzed by an enzyme called thymine dilate synthase, thymine dilate synthase. So that is how in biological cells the thymine is synthesized or uracil is synthesized okay. Now this is in biology now what about the real molecules, these molecules that have been evolved. So if you ask a question how come these molecules were evolved when there was no bio-molecules, when there was no biological cells when there was no biology, in fact, I am talking about Pre-biotic Chemistry.

(Refer Slide Time: 17:00)



Pre-biotic Chemistry prebiotic means before the biology when there was no living cells, there was no living thing in the world, in the earth. So we are talking about millions and millions of years before the earth was evolved before the bio-molecules was were evolved, before the living organisms were evolved. So how what was the starting point how the molecules were synthesized, from the materials that were there present before the biology.

There were no enzymes the in fact there were no bio-molecules there. And there were also no large larger molecules only the smaller molecules were present. So a popular theory is that before the earth I mean at the very beginning of the earth, that is millions and millions of years ago everything was in a very molten state, kind of liquid form that is being called as primordial soup.

If you have not heard this term, you will hear this term and quite often, if you try to read the evolution of the molecules or evolution of biological molecules on earth. So primordial soup is considered as the molten state at the beginning of the earth when earth was created because there was the temperature was very high there was lot of pressure thing and sparking were there. So, yeah and there were very less amount of air oxygen was very less so everything was in kind of an aerobic condition and this is the whole thing is called the primordial soup.

So we will now see how the some experiments have been performed to validate some theories that explains that will kind of give you ideas how the nucleobases were synthesized out of this primordial soup. So, when there was no biological cells when there was no protein, no DNA,

our basic question is, how the living cell was created? First so such a long kind of molecules and the proteins and DNA are very large molecules and therefore they are not easy to synthesize when there was nothing.

So obviously they had a starting point and again a popular theory and that is known as RNA world theory that predicts or that kind of gives an hypothesis that RNA was the molecule which was created first. Arena was the first bio-molecule which was created first out of all the other bio-molecules and from RNA the other bio-molecules were formed. And then came the single cell single living cell kind of amoeba kind of the living cells were created.

So now in order to synthesize the RNA you have to synthesize the nucleobases. So again it, it is it comes down to the fact that how the nuclear bases were synthesized at the beginning of the earth which is we call Pre-biotic synthesis. So now I will show you two beautiful experiments that have been performed. There are many people actually, there many researches are going on to understand or to predict this problem that how is or what exactly happened how the molecules were evolved.

And of course that has a big problem is that it has happened over millions and millions of years. So if you want to do experiments in the laboratory, you cannot spend that much time or you cannot afford that in that long time, to your reaction mixture to understand what could have happened. So you have to create a simulated kind of atmosphere in your laboratory to understand such factors. And therefore that makes it tough because the time is a big issue here.

The reason is when there was a when the small molecules were present in the primordial soup for millions and millions of years, they had high degrees of freedom. And of course over time I mean, if you think of that long time, then, the chance chances are pretty high that at one time or some other or the other time the molecules will come close together and they can do chemistry. And that timeframe is difficult to spend in the laboratory. So they have kind of simulated the atmosphere.

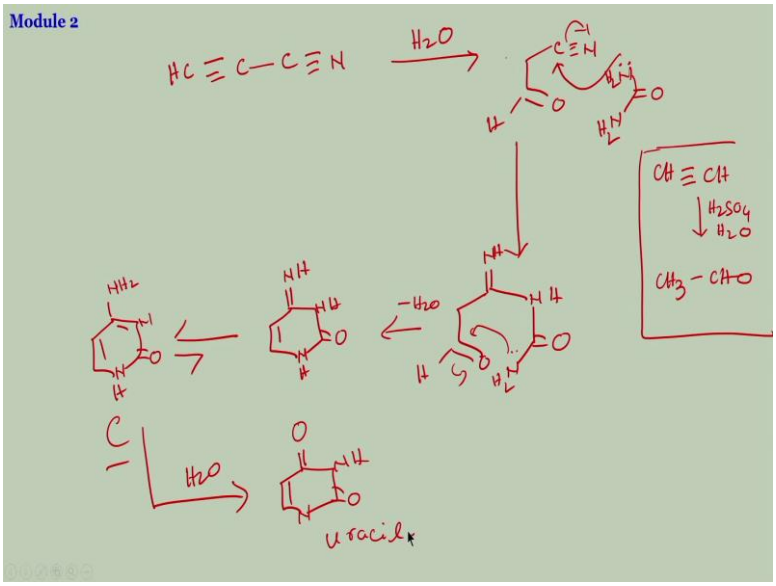
So first experiment I will show you is was published in Nature, the famous Journal in 1995.
This is the volume number you can search for it 375 and page number is 772. So that has actually tried to understand how the nucleobases were synthesized. And they have started with the simple molecules. So at the beginning of the earth all gaseous molecule on salts, liquid,

solid, everything was composed in a shrew and there are mostly a lot of gases were present obviously and very less oxygen as I am saying.

So it is very clear that the methane was present. Carbon dioxide and carbon monoxide both were present. So in this case they have taken methane and nitrogen. What they have seen, if you take methane and nitrogen, so this has been validated by other literatures also. And with high electric spark discharge you can be able to synthesize an organic compound, which is an acetylene derivative, acetylene and cyanide.

So this is cyanide this is a cytosine this is quite well studied actually. Many literatures have reported this reaction that happens. So this is a key compound now in our hand and from here I will show you how the nucleobases can be synthesized.

(Refer Slide Time: 23:54)



So if you start Cyanide and basically hydrolyze it with water, so it is like if you have the organic chemist can understand it very quickly. If you have a acetylene and if you treat it with H2SO4 in acid condition and water in acidic condition, then what you get is the hydrolyzed product that is CH3 CHO, an aldehyde. So, here if you treat it with water under certain experimental condition then, what you can have is this compound aldehyde.

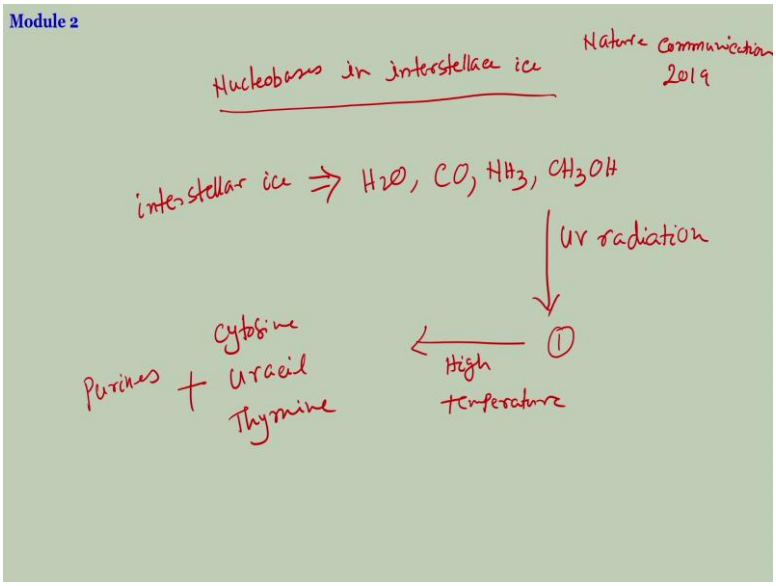
So this should be converted into aldehyde this would have been CH2 because free acetylene would be CH3 here. It is already bonded should be CH2 and you will have a cyanide as usual. And now if you react with Urea, Urea is a small molecule which is easier to synthesize or easier

to be evolved from ammonia. Ammonia and carbon dioxide or carbon monoxide actually produces urea that is also quite well known. So if you react with urea then what will happen?

Now you are close to the cytosine kind of synthesis. This so you will have N it will be H here NH of urea here NH2 aldehyde. Now, this can do the same reaction as you have seen before and I am skipping the step. So it will be minus water and you will have NH and obviously this is the immune form. So it will tautomerize at first, your cytosine is ready. So cytosine can be synthesized from this molecule. And of obviously this molecule came from only methane and nitrogen.

So once you synthesize hydrogen of course you can synthesize the uracil after hydrolysis this would be hydrolyzed or deaminated, the uracil was synthesized. So that is an experiment that actually kind of suggests or gives you hypothesis how the molecules the nuclear bases could have been evolved. There was one second experiment that I will talk about.

(Refer Slide Time: 27:20)



Another beautiful experiment actually and this was published in Nature Communication, another very high-profile journal in 2019. So what they have done is they have tried to synthesize the nucleaobases in interstellar ice. So the ice form in other stars or other stellar systems where there is no life yet can be a good mimic. Or can give you idea how the molecules could have been synthesized on earth because I mean, people kind of think that the same molten state or the same kind of molecules that are present in the interstellar ice form or in interstellar systems where present also on earth millions and millions of years ago when the biology was not there on earth.

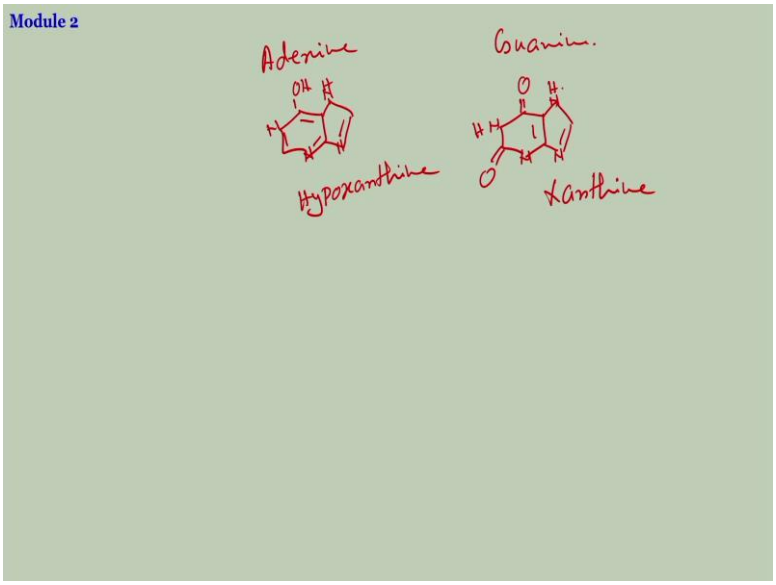
So studying the interstellar system and trying to find the evolution from there would have been a good mimic. So and again obviously you do not have interstellar ice in your hand in your laboratory. So you simulate in the laboratory what can the same composition that would have that are present or that have been already studied by the scientists that these kinds of molecules are present in the interstellar systems.

And interstellar ice is composed of water, carbon, monoxide, ammonia, methanol and all sorts of stuffs. So people have used the same composition in the laboratory and tried to find out whether you can synthesize the nucleobases. So what they have done in this paper is that they have taken all of these and of course at the beginning of the earth there are a lot of radiations. There are a lot of UV radiations intense UV radiations were there along with high heat, high temperature.

So they have used a high UV radiation in this composition. They have taken a composition of this all these molecules and they have started irradiating with high dose. Whatever they have obtained after that they have processed it further with high temperature because that was there in the prebiotic earth. And after this, they have tried to find out, what are the molecules that have been formed. And guess what they have isolated all three of the pyramidion eclipses.

They have isolated cytosine, they have isolated uracil, they have isolated thymine, all three of them were present in the final composition after this reaction.Not only that plus some of the purine nucleus derivatives like adenine guanine derivatives were also found in this experiment. Purines I will give you some structure.

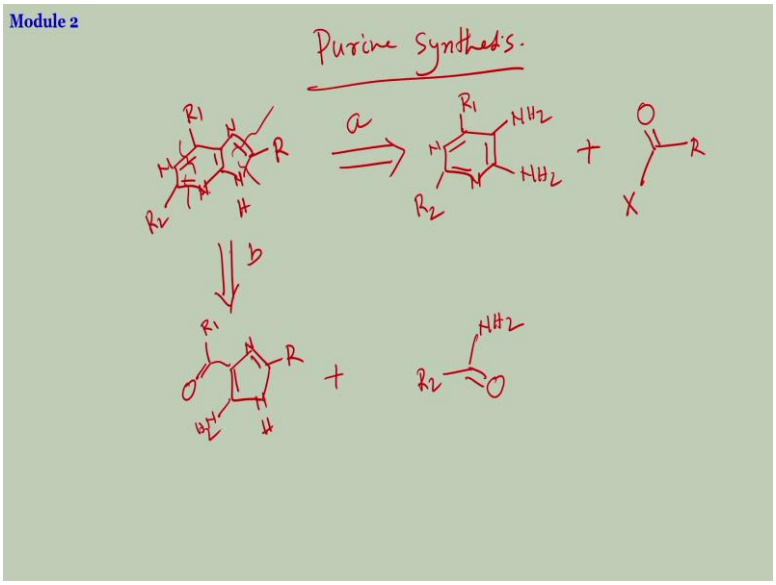
(Refer Slide Time: 31:17)



So adenine was there out of this purine bases adenine was found and then this compound should be 5 membered NH, OH, N double bond this is N here this is hypo xanthine, xanthine one kind of purine, they had also guanine, that was found and then xanthine was found. This is xanthine, this one was also found. So a mixture of the purine bases were also found in the same reaction mixture.

So this kind of chemistry actually tells you to develop more hypotheses or more predictions of about the fact that how the molecules could have been synthesized, how the basic subunits of the bio-molecules would have been synthesized from the pre-biotic earth, from the when there was no life. So similarly we will see that from some prebiotic chemistry later when we will see this amino acid synthesis. So now you have the pyrimidine kind of synthetic pyrimidine synthesis.

(Refer Slide Time: 33:24)



Now, come the Purine synthesis. So Pyrimidine's you can synthesize in the laboratory. Now comes the fact the how can you synthesize or how can you find the Retrosynthesis retro-synthetic routes for that Purine nucleobases? The general structure of the purine is, is this here the sugar would be attached, so this in nitrogen has to be free, has to be NH, here can be any derivative usually there is nothing here. But I'm keeping R so that you can have radiations.

Here there is N R2 depending upon the nucleus, they will vary. And N, N another in so if this is the structure of a purine how can you synthesize the target molecule? If you look at the molecule and if you think of the pyrimidine kind of synthesis, of course, you now know that this 8 member ring you can synthesize very easily. So one of the idea can be that you start with the 8 member ring and synthesize the 5 member ring on top of it.

In other words, you have to synthesize this 5-member ring on the pyrimidine. So now if that is the approach, I call it approach A or route A so then you have to think of how to synthesize the 5 membered ring. Now where to cleave the bond? As we have seen before that cleaving the carbon heteroatom bond is the easiest own is one of the easy tactics that can be used in organic methodology.

So in that case it could be that this bond can be clipped, this bond can be clipped and the whole thing can be attached to it. So if, if you do that then there is an issue that I have already talked in the last lecture. And then you have to fuse, you have to form the CN bond you have to form this CN bond also with an aromatic ring and now aromatic ring is already high in electrons electron rich.

So doing chemistry or making a new bond formation with the aromatic ring with another electro electron-rich Center is hard because nitrogen is electron rich, aromatic ring is also electron rich. So this one formation would be little bit harder possible there are a lot of methods are there this is little bit harder. So the easier way to find to do is not to touch this bond, not to touch this bond.

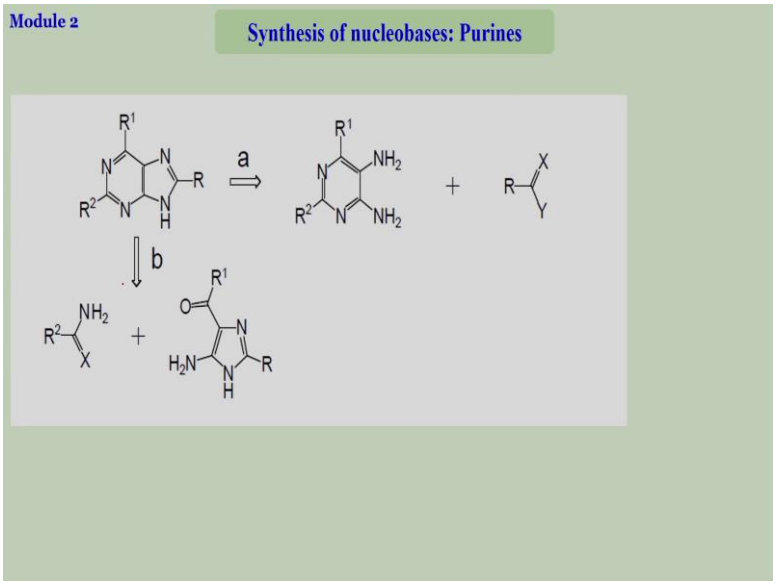
So but to clip this, if you clip this bond there so what you will get down to and if you clip this bond here, then, you can start with this amine here, this amine here and in this case what will be in the other compound? You are forming a double bond O, so obviously there has to be a double bond so because remember it was dehydrated and formed double bond O that could have been here and here it is a simple nucleophilic substitution.

So you can add just a leaving group good leaving group here in this case alright so this Plus this you can using these two you can manufacture the Purine. The other approach is you start with the 5 membered ring and then synthesize the or develop the 5 8 membered on top of it. So that is approach B. So you start with this and then develop the 8 membered again you can avoid the direct synthesis here because this is also aromatic.

So this point should not be clipped this one should not be clipped so you can chop of this or you can chop of that bond. So that will tell you have R1 and you have to form this double bond so it should be double bond carbonyl is a carbonyl compound. And in this case what can you have? You can have amine plus the R part R2, if this is amine and there is a double bond here so there has to be a carbonyl compound in this case.

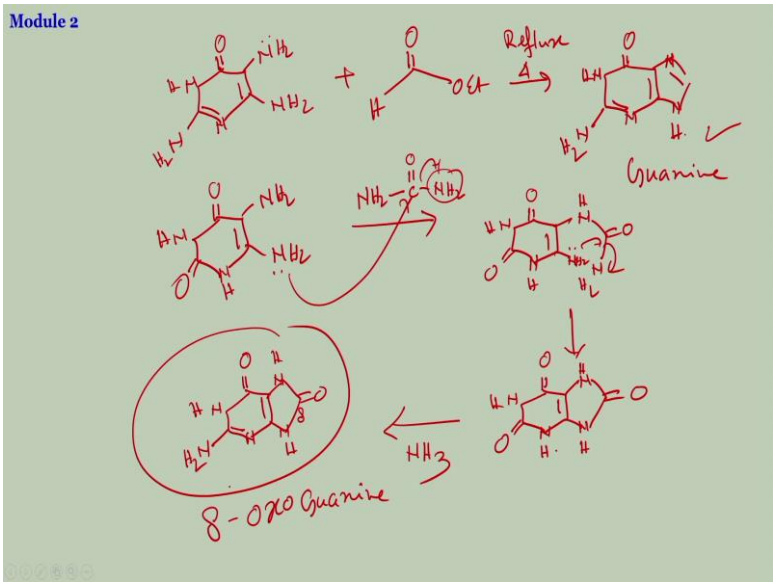
And this has to be manufactured, so there has to be an amine kind of this. So these two can give rise to the purine or these two 8-membered ring and this molecule can give rise to the same Purine. So these are the two approaches that are more viable so that is given here.

(Refer Slide Time: 39:01)



If this is the ring then you can clip path A and this is path B. Now how to do the forward reaction? So starting with this let us move to the forward direction.

(Refer Slide Time: 39:19)



If we take this molecule and I am starting one of the starting materials plus H then, what can you have do reflects almost the same condition and heat in the organic solvent then what we'll have is this one will react, it will come down there. So if you do the mechanism you will see that you can form nitrogen. Nitrogen NH here double bond OH will eliminate. So you will have amine, this compound. And this is your guanine.

Now you can see that there is another extra amine here which sometimes interfere with the reaction system. So in advanced chemistry you may need to protect this amine. Most of the times you do not require because this amine does not react that drastically so normally if you

use this you will get a certain amount of the guanine. Now if you think of little bit variation this if I start with and if you treat this with urea, can you find out what the compound going to be?

So in this case one of the amine would be a leaving group this would leave and then the dehydration will occur. What we will have is NH CO NH 2 and NH 2 here NH there was NH here this and then our second displacement of the amine group will give you a double bond over here. This would be NH, this should be NH, this compound if you do, if you treat this with ammonia then, this carbon will get emanated actually. It is quite easy reaction.

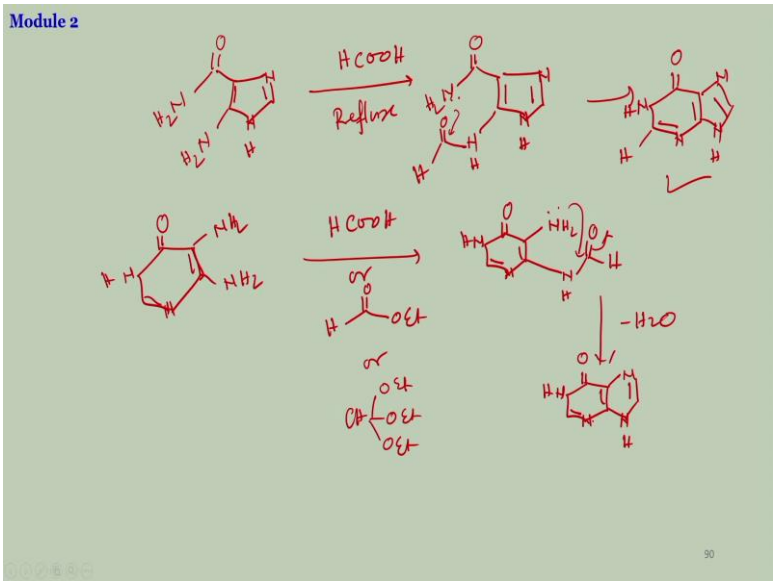
I will tell you why we are doing this. Then this would become amine and this would be double bond here. So you ideally could have started with this compound also here. And we could have gotten get down to the same this two would have been identical. If you would have started with this, with amine in the group already present. But in order to avoid the reaction of this amine, we have started with the double keto form here.

And then if you need an amine you can always do the amination by treatment with ammonia. So this compound this is a guanine derivative without the double bond oxygen here. This is this component particularly is known as 8-Oxo Guanine because this is the 8th position. Eighth position there is an Oxo group. So 8-Oxo Guanine is the one in this molecule is actually present it is a very notorious molecule which is created from the DNA.

Or the we call it the DN ellisons, the damage of DNA creates this model, this molecule. This molecule is the reason why certain cancers are formed. It induces the mutations and that we will see later. 8-oxoguanine is formed by oxidation of guanine that is present in DNA. And this under creates the mutations that causes the cancers, cancers a lot of different kinds of cancers and such as lung cancers, pancreatic cancers.

They have they are because of this formation of this molecule 8-oxoguanine so which you can actually synthesized in the laboratory in this way.

(Refer Slide Time: 45:09)



So now the second if you start with the 5 membered ring how can we develop the 8 membered NH here and if you start with the amine, this instead of this amine, it can be any other living group also you can start with an ester as well or acid chloride as well. If you treat it with formic acid, reflux I am giving just one example. You can change you can by varying this reagent you can make different compounds, different molecules.

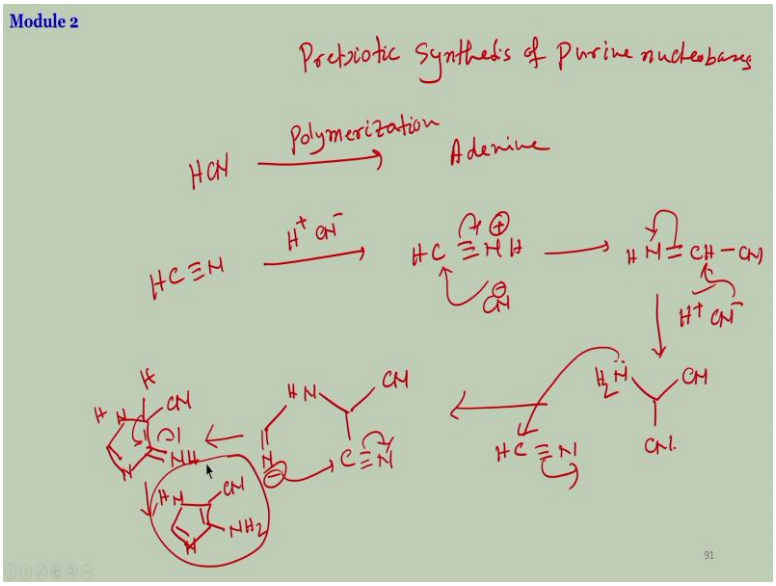
Then what this will do is this Amine would react and would give you in NH CHO. And of course now this reaction can take place. So you will have the 8 membered ring and then the 5 membered N here N here and then it can be an NH, it can be H here initially there will double one here. There now it can actually tautomerize you can write it in in any form you wish. This so this should be formed initially there will be double bond here and it will be tautomerize into this.

So that is the way you can produce the you can start with the 5 membered ring and can produce the 8 membered compound. So this is your basically Adenine compound. Now the same adenine so you have synthesized it from the 5 membered ring if you want to synthesize the same compound from the 8 membered then you can start with this here the amine, amine nitrogen NH. So this compound you can treat it with the same formic acid or as you have used this format, L format or vt ortho format.

Any one of them you use you get down to the same compound that is an aldehyde fused with the amine group. They should be produced and of course then followed by dehydration minus water. Double bond here, this would be an NH, there was the double bond present here, this if

you go back to the pre-biotic, I will give you one example the how the purine nucleus is obviously you have seen one, that the mixture of ammonia, carbon monoxide, on all these things have produced the pyrimidines as well as the purines.

(Refer Slide Time: 49:20)



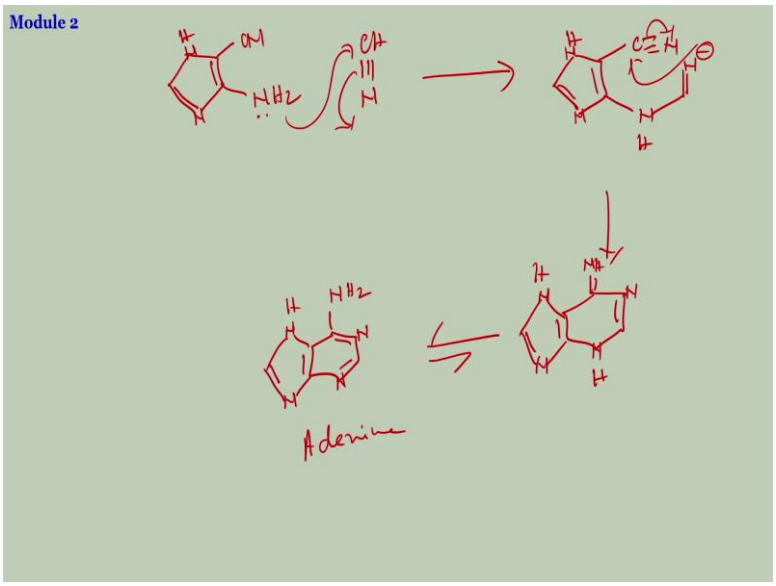
There are other methods that have explained pre-biotic synthesis of Purine. Purine nuclear bases basically and this one is actually very interesting. It all started with the HCN hydrogen cyanide and it's basically a polymerization of HCN means that many of HCN molecules were added together one after the other and form the final compound. We will see how adenine was developed. So first if you have HCN and treat it with another molecule of HCN, so HCN is basically an acid.

So this is H + CN- I am writing so it can be protonated and then so if this is protonated then what you get is HCN + and then CN- would react here eliminating this, so you will have in this case, if you are write NH double bond CH and Cyanide, this molecule. Now this, can further be protonated and will produce so this will further protonate it will become NH 2 plus this will move and there were another CN- would bow basically this form.

So this would be NH 2 this is cyanide they would be another cyanide here. Now once you form these, now you have a nucleophile amine which has a lone pair of electron. So cyanide is also electron deficient it acts in both ways. So this lone pair of electron can attack the cyanide carbon and eliminate this. So in this case what you will find is cyanide, this NH, H, this is H double bond N, double bond N N- and there is one cyanide here.

Now this minus can react here eliminate this. So now you have a 5-member ring within there is a double bond here carbon, then you have, NH cyanide and in this case, you have double bond NH. Now this double bond NH would obviously become an NH2 by tautomerism. So this would be NH2 here. So here there is one H, so this can tautomerize NH. So you will form this compound I have to write again here.

(Refer Slide Time: 53:20)



In cyanide in NH2 there iss a double bond here and then an NH so this react with another HCN, another HCN, so obviously here and this will come down in NHC double bond N minus rest is all the same NH. Now this can go here and if you do the whole thing then you will see that you have so this will be double bond NH and this is your N double bond N here NH now this will all tautomerize because this has to be NH2.

NH this, this will become NH2 and this will become N this will be another double bond here, so, which is your adenine. So that is how you can synthesize whole of the purine nucleus adenine out of the HCN cyanide. There are other purine nucleus also where explained how you can synthesize them from HCNs. So I mean this is one just one example of showing how the pre-biotic chemistry could have taken place for the synthesis of the purine nucleaobases.