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

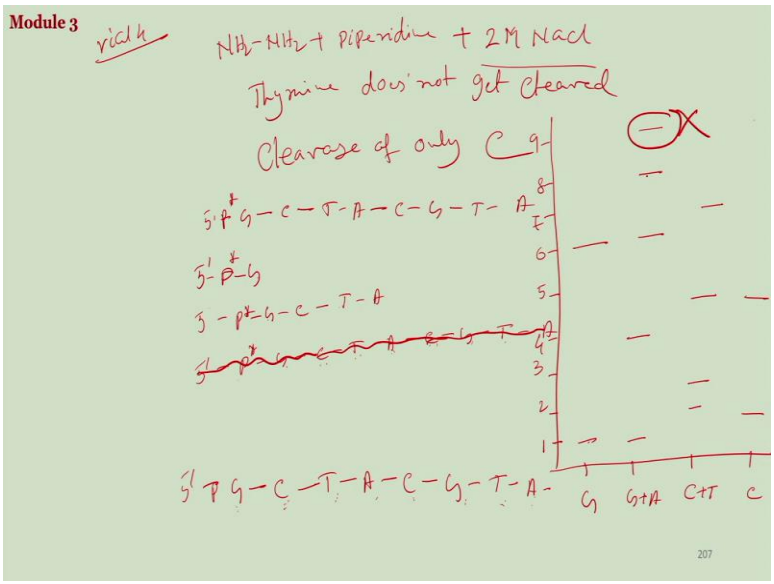
Lecture No. 15
Numerical Problem – 1

Hello everybody and welcome back. So far, we have completed 3 modules. Module 1 was about the structures and characteristics, the chemical structures characteristics, some other properties of the nucleic acids, proteins, peptides as well as the lipids. In Module 2, we have done some organic synthesis, the laboratory methods or the industrial methods for the synthesis of the nucleic acids and their derivatives.

On the other hand, in Module 3, which was the last previous module just to have completed that was about the biology and doing in vitro studies are doing in vitro methodologies on DNA replication and how can we amplify DNA using the PCR polymerase chain reactions? And then finally, how can you sequence a DNA that was about we have talk about 2 sequencing methods, 1 was the max Sangers methods and the other one was Maxam Gilbert method.

So, so far we have completed these 3 modules I think we have covered quite a good part of the chemistry as well as some biology. And so far we have not done any problems solutions. So, today I thought that I will talk about or we will discuss about some problems numerical problems and then followed by later on synthetic problems. And then the solutions also you will discuss today.

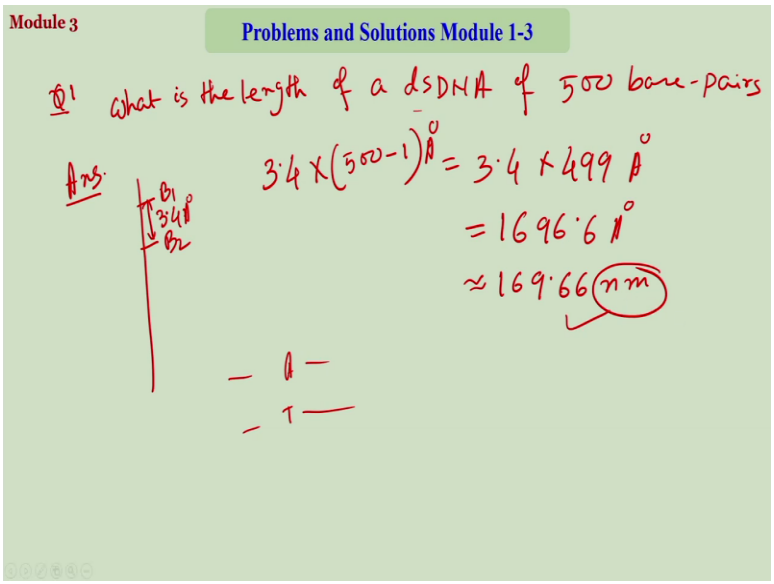
(Refer Slide Time: 02:35)



But, before that, I would like to clarify something which was the last lecture in the Maxam Gilbert method, the diagram, there was some mistake in it. So, I would like to clarify today. Now, that our target I have, I think, I have found out the issue or target DNA was this was our target sequence, which was 12345678 total 8 number of nuclear basis and here we have done 9. So, that was the problem, this one should not be there were until 8.

And this issue came because I had written the whole the full length o f DNA, which is not a cleavage part, this is present in all the other vials. So, this will not come up for consideration. So, as you see the fragmented patterns are stopped here in vial 4, and then we are good, then we are good everything else is fine. So, only thing is that since we have 8 nucleobases long DNA, this 9 will not come.

(Refer Slide Time: 03:38)



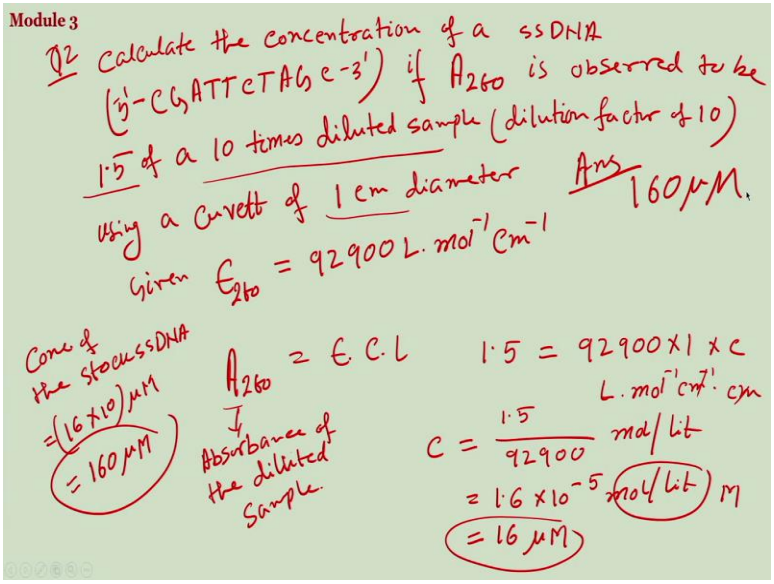
So, problems and their solutions I will discuss about quite easy problems actually not very difficult to do. And some of them might appear in assignments as well. Not quite the same but close, closely related. So, first one question number 1 is to start with a very easy one, what is the length of a double stranded DNA, ds DNA of 500 base pairs. So, if a double stranded DNA has 500 base pairs in it, which means the DNA is 500 more long, then what will be the length of that DNA?

Of course, this is assuming that DNA is straight in nature or the double helix in nature. So, this is a very easy problem but actually you can solve it very quickly I guess. So, the difference or the distance between 2 nuclear bases is in a given DNA. If this is a DNA, I am just writing this rate. If this is base 1 if this is base 2 then this distance is as we know is around 3.4 angstrom. So, therefore, your answer will be 3.4 multiplied by your 500 nucleobases long which means 500 - 1 if you want to really accurate that we need number of distances will be there.

So, this is about in angstrom 3.4 into 499 angstrom or this will come as 1696.6 angstrom that will be length of the DNA which is almost close to 169.66 if you want to convert it to a different unit nanometer we can convert it to any other unit also micro meter like that so, that is how of course, you can calculate the length of a DNA. Of course, if you say that length of a single stranded DNA having 500 basis will be the same as this.

So, base pairs, we generally use the term base pairs to mean that you have 1 base and alternate to the other base. So we call this as a base pairs and this is the usual terminology which is used when defining a double stranded DNA or when defining a single stranded DNA, we call it we just call basis. So this is your first problem.

(Refer Slide Time: 07:04)



Question number 2 calculate the concentration of a single stranded DNA, which has the sequence, 5 prime CGATTCTAGC and 3 prime, this is the sequence of the DNA if absorbance at 260 is observed to be 1.5 of a 10 times diluted sample or you can tell these as with the dilution factor of 10 using a cuvette of 1 centimeter with give it up 1 centimeter diameter given if molar extinction coefficient at 260 is close to 9 to 900 liter mole inverse centimeter inverse.

So this is the question that you have you are given a single stranded DNA which has the sequence and when you are measuring the EV absorbance at 260 nanometer you have used a 10 times diluted sample from the parent DNA stock you have from the stock solution from the mothers single standard DNA solution, you have done a 10 times dilution and that 10 times diluted sample was used to observe the EV absorbance and that EV absorbance is 1.5, the value of 1.5.

And you are using the cuvette which is 1 centimeter basically, the epsilon value is given so, question is what is the concentration of the mother single stranded DNA. So, basically you have

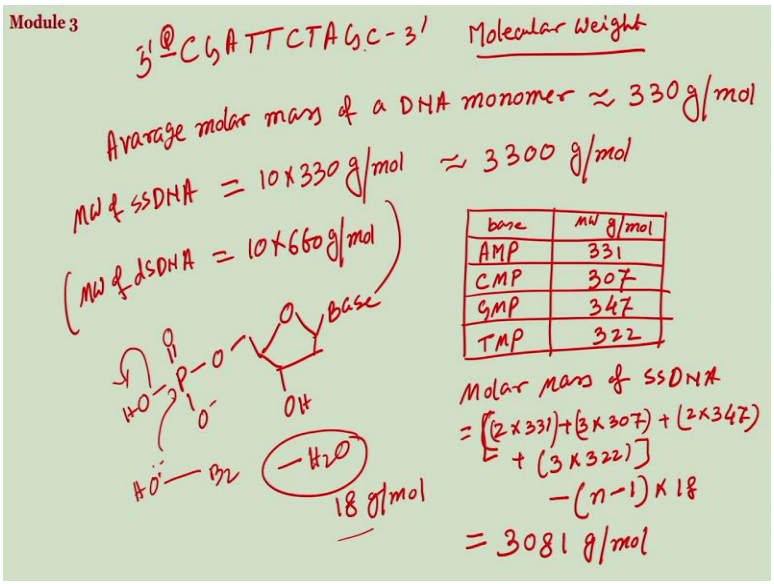
to calculate the concentration and in order to calculate the concentration, what is the basic equation, Beer Lambert law that is observance at 260 = epsilon C L. Epsilon is the molar extinction coefficient, molar extinction coefficient which is given here, see you need to calculate L is the path length or the length through which the light travels.

And this is given us 1 centimeter it is basically the cuvette you were using that same diameter of the cuvette will be your path length because that is the length or that is the distance that will be covered by the light. So and this A value is given us 1.5. So 1.5 = 92900 into L is 1 into C concentration, what is the unit liter mole inverse centimeter inverse into the length was in centimeter. So this cancelled out.

So, your concentration will be 1.5 divided by 92900 mole per liter and that is coming as 1.6 into 10 to the power - 5 mole per liter which is equivalent to 16 micro molar. So, moles per liter is molar capital M micro molar is 10 to the power - 6. So this would be 16 micro molar and that is the concentration of that 10 times diluted sample. So this is basically is the observance of the diluted sample.

So therefore concentration of the parent DNA stock I am writing stock single stranded DNA that would be 16 micro molar into 10 times or 160 micro molar. This should be the answer to this question. So, unsure basically is 160 micro molar.

(Refer Slide Time: 14:31)



So, very quickly we have used this sequence, CGATTCTAGC and 3 prime if I ask you to calculate a molecular rate of a DNA how we are going to do sometimes you have a very long DNA, of course. And usually you do not have to do anything; you can just type the sequence somewhere. There are many places in the internet, or there are excel files also available, where you can just type your sequence and it will give you the molecular weight.

So, ideally, you do not have to calculate it by hand. What I would like to show you or I would like to discuss how you can calculate it if you really want to or if you need to, it is good to know. So, what usually is a practice that when you have a long DNA, it is assumed that they are composed of all 4 kinds of nuclear bases? And in general, people think or in general it is considered that the average molar mass of a DNA I will write monomer.

So, we are calculating the molecular weight of this DNA strand. So, the very quick calculation is that it is considered that the average molar mass of a DNA monomer is approximately 330 gram per mole that is the number of which has been calculated taking the average of all 4 nuclear bases and it is assumed that your DNA will be composed of almost equal number of all 4. So, it is of course not the accurate measurement.

But an approximate measurement approximate calculation 330 gram per mole is the average molar was of a DNA monomer that includes the phosphates that includes the sugar. So,

therefore, for this single stranded DNA so, molecular weight of single stranded DNA this particular DNA will be how many bases you have 12345678910, 10 into 330 gram per mole. So, approximately 3300 will be your approximate molecular rate of your given DNA. So, this is for the single stranded DNA.

If you have to calculate the molecular rate of the double stranded DNA if this DNA comes in a double strand form then it will be simply the double of it so, 10 into 660 gram per mole that would be the molecular weight of the double stranded DNA. This is for very quick calculation and very, approximate calculation. This is not at all close to the perfection of course. Now if you really want to calculate the individual taking into consideration about the individual nucleobases, then you can do that too.

So I will give you a chart which you can use later. So, this is the base, the nucleobase and this is the molecular weight in gram per mole but if you have adenine, adenosine AMP adenosine monophosphate then its molecular weight 331 if you have CMP cytosine monophosphate or cytidine mono phosphate that has the number 307. GMP is 347, TMP is 322 gram per mole and this the structure is of course, this wage.

Here is the base and this is mono phosphate means O - H basically now, if you want to calculate in this table, if you want to calculate the molecular weight of the given DNA, then what do you have to do? Basically use sum of all these molecular weights of the individual basis. But that only accurate because when you are doing the phosphodiester bond formation which of the other base B 2 is coming and they are doing this reaction that is why you are getting the phosphate bonds.

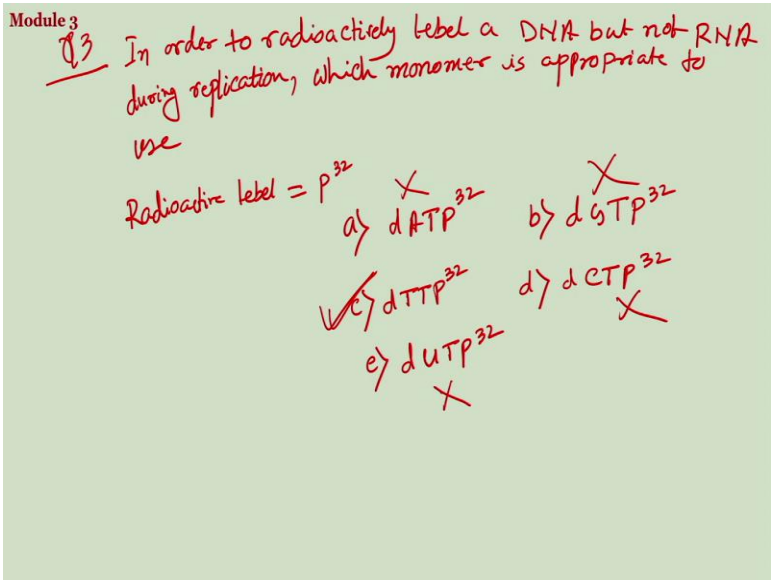
So it is a minus water molecule is released, 1 water molecule is released which has the molecular rate of 18. So, you hope to sum of all the individual nucleobases in terms of the mono phosphates and then minus that a number of phosphate bonds that are there multiplied by the 18 that will give you the correct result. So, in this case, if I calculate molar mass of your single stranded DNA that would be so 2 into how many how many adenine is there 2 into 331 plus cytosine 123,

3 into 307 + G, 12G into 347 + T 123 T 3 into 320 minus how many phosphate bonds number total number of basis minus 1. 12345678910 and there was these.

So, in general formulize n -1 into 18 number of basis minus 1. So, if you calculate in this case. So if you take that into consideration your total value will come as 3081 gram per mole. So, here you have the phosphate. So here nothing is negative - 123456789. So this should be 9 into 18. So 10 – 1, n - 1 is your now 10 - 1, so 9 into 18. So this would be your, the calculated molecular weight, but only thing is that you have to remember this step.

Similarly there is you are seeing the epsilon and the value that now nowadays we just if you just type the sequence it will give you the value of the epsilon of the molecule. But you can also there is a way to calculate also, It is little bit complicated and to some extent is boring. That is why I am not sure what you can maybe if you are interested you can go through it, how to calculate the molar extinction coefficient of DNA.

(Refer Slide Time: 25:09)



So now question number 3 in order to do radioactively lebel a DNA but not RNA during replication then the question is which monomer is appropriate to use. So, suppose you have a mixture of DNA and RNA in your vial and there is replication going on now, you want to track only the DNA and therefore, you want to level only the DNA with the radioactive P 32. So, radioactive lebel means we are labeling, we are using P 32.

So you want to do it selectively for the DNA and not the RNA. Then the question is, I will give you the options. You have to pick up the right kind of monomer that you will use in order to get the levels DNA. Option a is dATP with the level 32. Option b is dGTP, option c is dTTP 32 level and d is dCTP this last one, dUTP. Now out of these 4, which monomer you will actually use to tag the DNA selectively and not the RNA. So, of course it is easy.

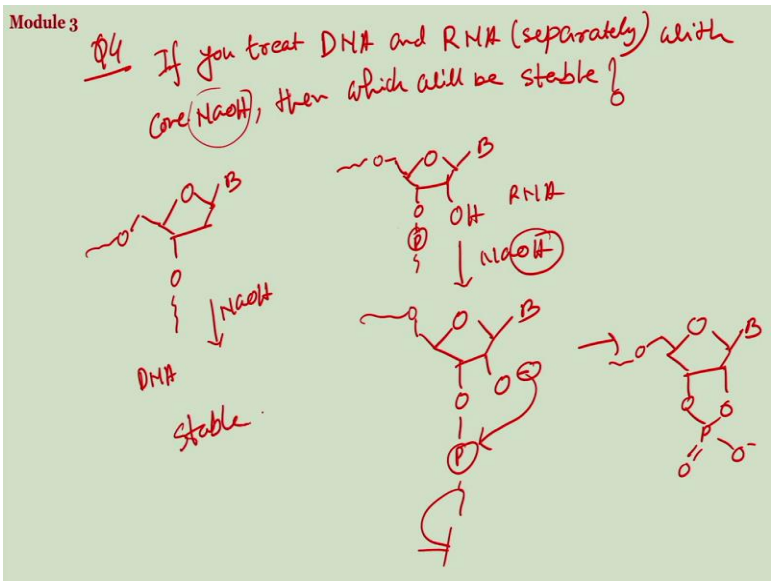
If you just think of replication, now when there is replication going on, what you need you need polymerase, you need the primer which is there and you need the free dNTP is should be picked up and the extension will happen. Your DNA is their target DNA is their target RNA is their both are having amplification, which essential means that you have a DNA polymerase you have the RNA polymerase also and all this dNTPs.

Now, you find out the monomer which is exclusively used for DNA and not RNA. So, if you start from the top ATP, ATP is a monomer for DNA of course, DNA has a denotion RNA also has that denotion. So, ATP would be taken up by both DNA and RNA during the replication. So, this cannot be it will be common, if ATP is taken up, then both DNA would be level RNA would also be level GTP warning is present in both DNA and RNA.

So, similarly, if this is if you use this then both will be level DNA plus RNA. So, this is not a selective 1 thymine TTP thymine is present in DNA, but not present in RNA. So, if you use this only DNA replication during DNA replication, this will be picked up RNA for RNA replication, this will not be chosen because it is not present it is not the monomer for RNA. So, therefore, this should be the correct solution TTP with the labeled phosphate you should use if you want to selectively level the DNA CTP is present cytosine is present in both DNA and RNA.

So, this cannot be used is presenting RNA not in DNA. So, if you use these this will label the RNA and not the DNA. So, this is also not correct. So, you hope to use the dTTP then then you can label only the DNA target or only the amplified DNA sample specifically or selectively.

(Refer Slide Time: 31:22)



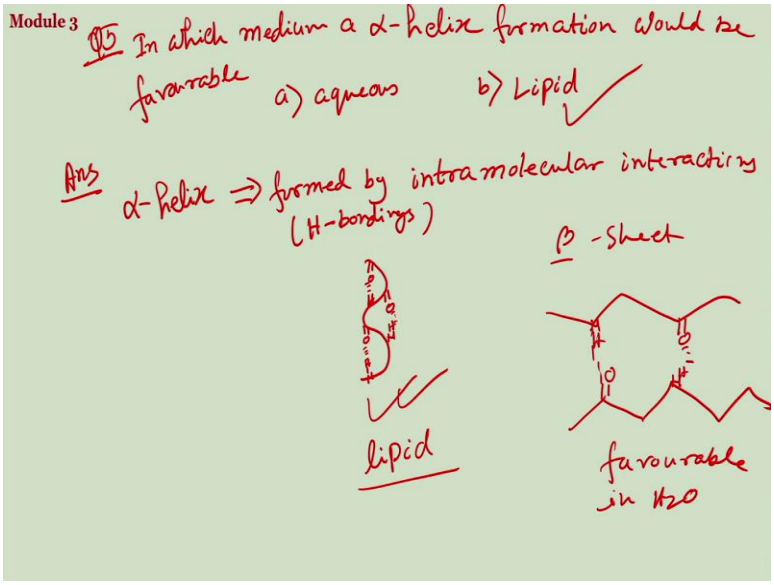
Question 4 is you have DNA and RNA. Question is, if you treat DNA and RNA separately with concentrated sodium hydroxide then which will be stable is DNA be stable in sodium hydroxide or RNA will be stable in sodium hydroxide or none of them will be stable in sodium hydroxide. This is base so this is your DNA without the hydroxyl would I think the phosphate P and then this is the RNA with the OH here.

This is your RNA, if you treat this with any which what will happen? The fact is RNA will be degraded or RNA is not stable in sodium hydroxide. On the other hand, DNA is somewhat stable in sodium hydroxide. The reason is you are using a strong base, so hydroxide is a strong base and that will deprotonate into minus that you do not have here. So, phosphate is there the rest of the particle. Now, what this is going to do this becomes a very good nucleophile.

Now, very reactive nucleophile also, this will immediately react, this reacts here and it basically the other phosphodiester bond. So, rest of the part of the RNA would be cleaved off and the first thing that will get is phosphate and there is nothing else here as usual, you have the other part, then followed by the other reactions but initially, you are here your extension the other part of the RNA will be chopped off and that will be happening for every phosphate group for every phosphodiester bond that you have.

So, every nuclear side positions or nucleotide position would be for the backbone would be broken up or be destroyed in RNA, if you use sodium hydroxide, because that is a very good base. So, RNA will be degraded or not stable, DNA would be stable.

(Refer Slide Time: 35:39)



Question 5 in which medium this is about the protein structure alpha helix formation would be favourable in which medium alpha helix formation would be favourable alpha helix means alpha helix of the protein a is aqueous or b is lipid formation of alpha helix would be favourable in aqueous medium or in liquid medium. This is the question what is the answer? So, when we look talking about alpha helix how is it formed? Alpha helix is formed by intramolecular interactions.

This is formed by intra molecular interactions primarily the intra molecular hydrogen bondings and intra molecular hydrogen bonding basically means the bond between the backbone carbonyl and backbone NH. This is the hydrogen bonding there can be like this I am just writing NH here in NH in the backbone they will form hydrogen bonding and it is because of this hydrogen bonding that the strand to go to the helical confirmation.

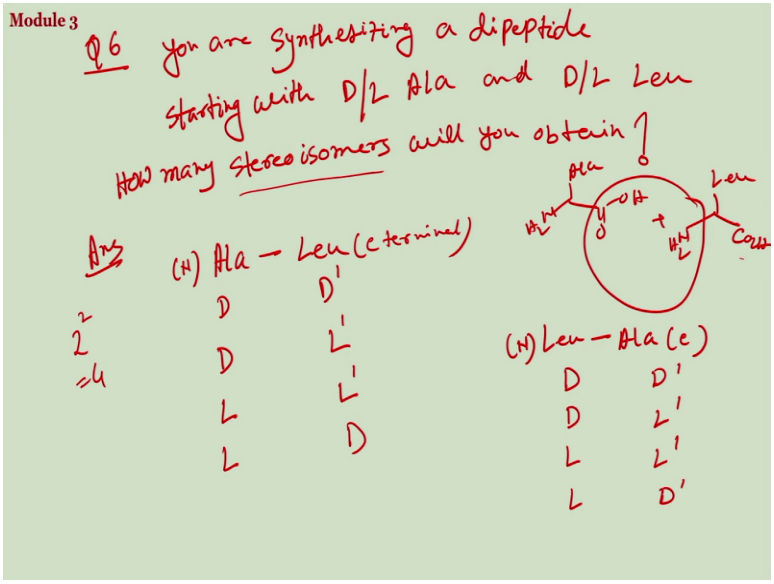
Because that way the backbone carbonyl and the backbone NH can come into the bonding distance that is the driving force for the formation of alpha helix. Now, the question is in which medium this would be favourable aqueous or lipid. So, which is the other form is the beta if the beta confirmation is formed, then that is because of as we have mentioned intermolecular

hydrogen bonding in NH and CO is 2 molecules CO and NH like this that will allow the formation of a planar structure.

So, within a molecule, what will happen? If you keep it in aqueous medium, then the chances is high that your polar groups in NH and CO both are polar, they would be one to be exposed towards the aqueous medium that means towards the outside of, the backbone and that will increase your chance of beta sheet formation. On the other hand, if you put it in a lipid is insoluble, lipid does not have water liquid is hydrophobic.

So, the polar groups would want to hide away from the surrounding medium. And how can this hide away this is one good way that they can have intramolecular hydrogen bonding forming a kind of structure, which will be away from the outside environment. So, chances of getting the intramolecular interaction would be better in the hydrophobic medium, in this case the lipid so, alpha is formation will be more favourable in the hydrophobic medium. Just a reverse for beta sheet it this will be favourable, in water medium aqueous medium or hydrophilic solvent.

(Refer Slide Time: 40:45)



Question 6 another question about peptides is suppose you are synthesizing we will talk about the synthesis of the peptides later. But this is a very common thing. You are synthesizing a dipeptide starting with this is mainly for organic chemists, D and L a mixture of DNA basically,

you usually the common alanine is will be L conformations in protein. Here we are using both D and L mixture of D and L alanine ALA for alanine and D and L of leucine.

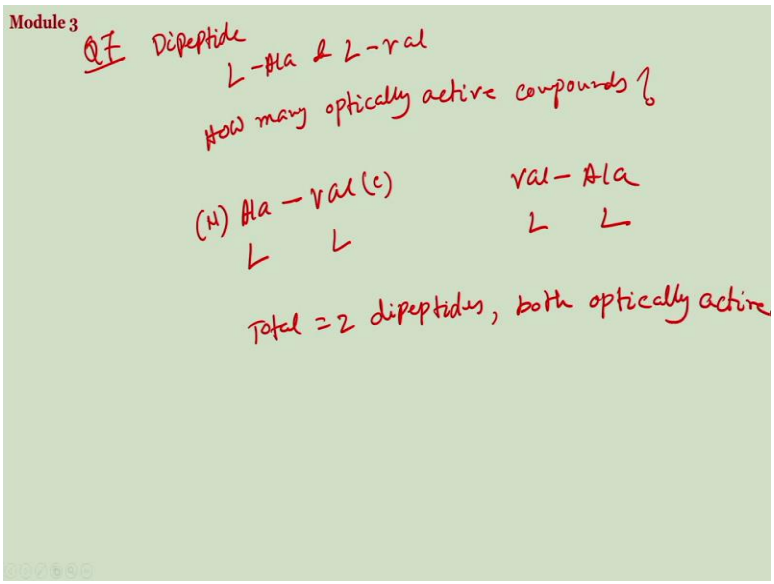
Then the question is how many stereo isomers will you obtain? That is a question if you form a dipeptide between alanine and leucine and both of them have the 2 optical isomers then total how many number of stereoisomers that will be heavy. So, if there is a dipeptide between alanine and leucine and how many dipeptides can be formed. So, these can be the N terminal alanine reacted with the C term reacted with the leucine and having a C terminal end free.

This is N terminal this is C terminal, which means very quickly I am just writing alanine here which just writing leucine, because alanine and leucine, none of them have any functional group. This is acid so, if you react here then you have the amine free for the alanine which is called N terminal and your acid will be free for leucine which is called the C terminal. Now, it can be the other way around also.

So, one is this other one is leucine and alanine which means the N terminal in the leucine and C terminal in the alanine. So, these 2 are the dipeptide kinds that will be in your solution in our product. Now, you have to look at this stereochemistry you started with D and L both also for leucine has started with D and L both so if this is D, this can also be D, writing the prime. This can be D, this can be L. This can be L is the same thing.

No, is not the same thing. L this can be L here on the prime. This can be L this can be D 4 same here D D prime, D L prime, LL Prime, L and D prime. And all of them will be to chemically active because these are 2 different molecules. So, if you remember your basic organic chemistry, if you have 2 optical active centers, the total number of stereoisomers that is possible is 2 to the power 4. So, 4 for each 4 for here 4 for their so, total 8 number of stereoisomers that will be possible for dipeptide.

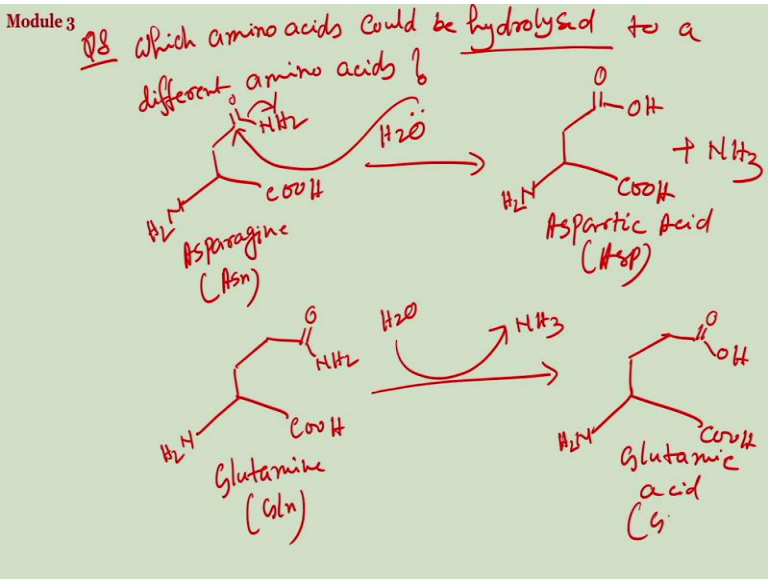
(Refer Slide Time: 45:51)



Now, another very quick question is you are doing the same dipeptide I will not right the whole question you are doing the same dipeptide but with L alanine and with L valine then was nice how many particularly active compounds will be formed? So very simple answer alanine so here you have only one steroidsation of alanine and only one steroidsation of valine. So you are dipeptide can be in alanine with C valine or you can simply write alanine valine and other case you can write valine alanine.

This is confirmation is locked or the stereochemistry is locked L this can only be L DD not possible because you have not use them here also L L so, how many optical compounds total they are really only 2 components are 2 dipeptides both optically active. This is very easy.

(Refer Slide Time: 47:41)

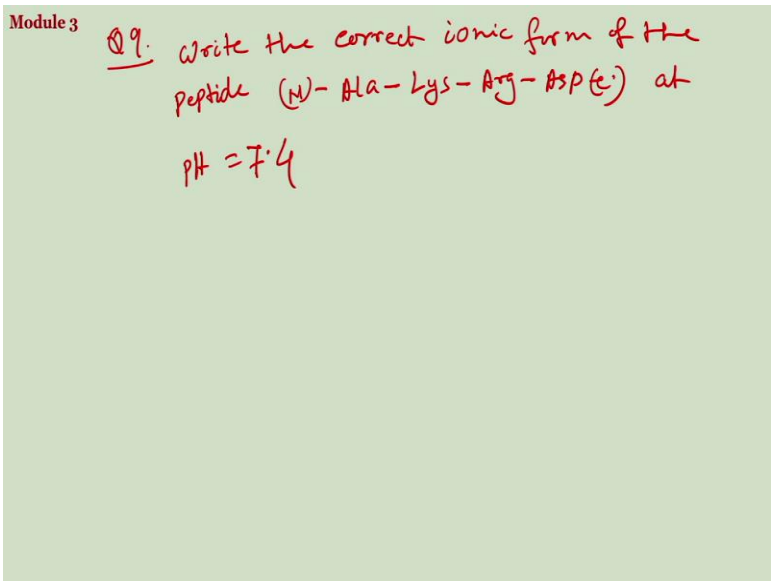


So question number 8 very interesting question actually which amino acids could be hydrolyzed to a different amino acids it is so if you look at the structure of the amino acids, from there you look at compounds or you to pick up the amino acids that you can hydrolyze in water and can get another type of amino acids. Can you think of it? This is your amino acid. So there are 2 actually often such type CO NH2 this is amino acids this is your asparagine in short ASN.

So, amino acids which have amide side chain which have a functional group or amide functional group in the side chain and amide can be hydrolyzed into acids if you hydrolyze it in water or if you use a little bit of proper condition optimized condition. So, basically aqueous medium is a hydrolysis then what do you get is water will attack here. These will both this will come back and eventually the amino will be eliminated.

So, here you will get OH COOH this one will be their plus you will release ammonia. So, elimination of ammonia would be there and this amide group will be replaced or would be hydrolyzed into a carboxylic acid group. Now, this becomes another amino acid this is your aspartic acid ASP that is one example. Second example is also having the same amide bond 2 CH2 and then this is your glutamine Gln is the abbreviation. If you take it with water and ammonia you get another acid or oxalic acid here as usual this is your another amino acid called glutamic acid Glu so I will write another question.

(Refer Slide Time: 51:53)



And I will leave you with this write the correct ionic form of the peptide which is given as in terminal alanine, lysine, arginine, aspartic acid, this is the C terminal. So, what will be the correct ionic form at pH of 7.4 physiological pH. What should we the proper structure of this peptide? We can do then I will show you the answer in the next lecture Thank you.