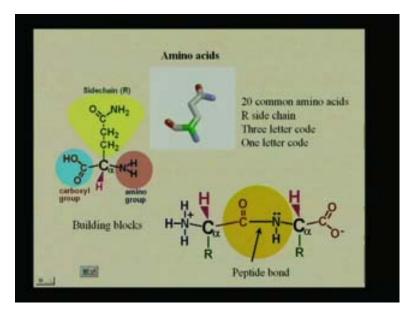
Biochemistry - I Prof. S. Dasgupa Department of Chemistry Indian Institute of Technology, Kharagpur Lecture –2 Amino Acids II

Welcome, we start with the discussion on Amino acids. As we have seen in the last lecture the Amino acids rather have an asymmetric carbon atom. This carbon atom is attached to an amino group and a carboxylic acid group. These amino acids have different side chains that we have already seen in the last lecture.

(Refer Slide Time 00:47 min)



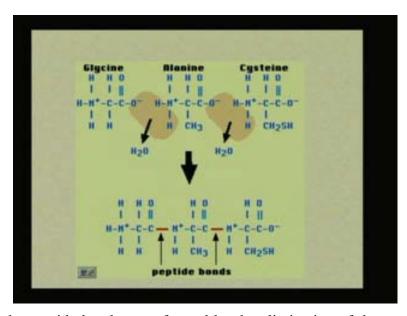
These side chains differ greatly in the property that they have. And they have associated with them by a 3-letter code and a 1-letter code. If we go to this amino acid this side chain correspond to the Glutamine here the 3-letter code for Glutamine is Gln and the 1-letter code is Q (Refer Slide Time 1:48 min).

Now if you notice in this stick structure we do not see the hydrogen atoms. Here the nitrogen atoms are always marked in blue, the oxygen atoms are marked in red and the carbon atom which is marked in green is for the asymmetric carbon and the others are marked in grey. This is the universal principle that the nitrogen atoms are always in blue, oxygen atoms are always in red and carbon atoms are in grey.

The next thing we looked at was the joining of two amino acids. The joining of two amino acids gives rise to the peptide bond. These peptide bonds have associated the carboxylic acid of the previous amino acid with the amino group of the next amino acid. So this R group corresponds to the first amino acid of this particular dipeptide and this R group corresponds to the second amino acid of the dipeptide which

means that in a polypeptide chain the first amino acid will always begin with an amino terminus. So it is never written any other way the first amino acid always have the amino terminus in a polypeptide chain, it always is an NH_2 terminus. And a carboxylic terminus is always the end of the protein chain.

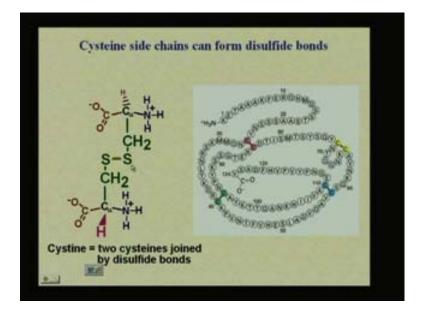
(Refer Slide Time 3:49)



As we know earlier the peptide bonds were formed by the elimination of the water molecules. In this case we have a try peptide and from the R groups we can identify the peptide. Now you can understand that as we have a protein chain or a poly peptide chain this is going to get longer and longer. In this try peptide chain the first amino acid is Glycine because it has two hydrogen atoms, this is Alanine because it has just the -CH₃ group attached to it and this is Cysteine because it has the -SH group attached to it. (Refer Slide Time 4:32 min)

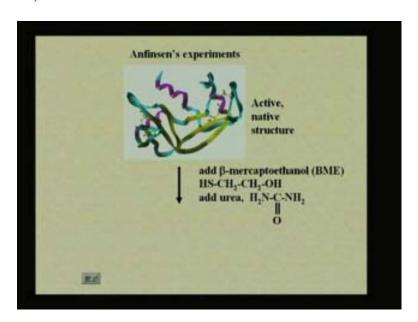
We also found out that this Cysteine amino acid can form a disulfide linkage. The peptide bonds are formed by an elimination of water molecules. So the carboxylic group of one amino acid looses its -OH with the -H of the amino group of the next amino acid to form a peptide bond and it is a strong covalent bond which is formed. The only other covalent bond is formed in proteins is the disulfide linkage. The disulphide linkage is formed by linking of two Cysteine residues irrespective of their distance. This is the SS linkage, this SS linkage that we showed here.

(Refer Slide Time 5:33 min)



Now I mentioned about this experiment where the whole protein chain was unfolded and it was found that exactly the same disulfide linkages are formed. When we look at a protein structure here we shown are an active native structure. What is an active structure means? It means this protein of this enzyme has its function induct. So when we looking at this active enzyme we know that its enzymatic activity is to the fullest because it is a folded protein it is in its native structure.

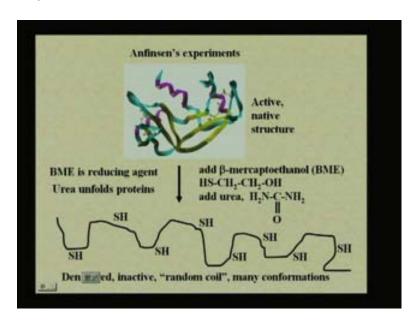
(Refer Slide Time 6:29 min)



Now, if you add β -mercaptoethanol to this it will reduce the SS linkages to form SH linkages. So now the Cysteine SS disulfide bond is a Cysteine residue over again. We have added β -mercaptoethanol which is denoted as the BME then we add urea. Here urea unfolds the protein which means that all the

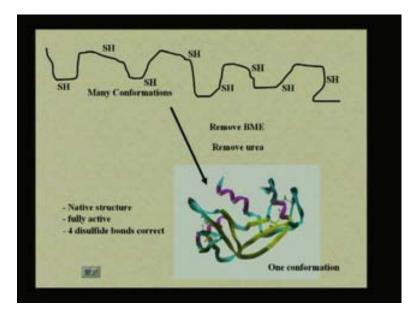
interactions that occur in the folded protein. The whole thing is unfolded to form a long polypeptide chain. Because we have the native structure and we have four SS linkages. What happen to those four SS linkages? We reduce the SS linkages to form SH linkages. So we added urea after forming the SH linkages. Once adding urea it unfolds the protein completely. So now we have this long chain with eight SH bounds with it. So it is a Denatured, Inactive, 'random coil' which means it is not folded any more.

(Refer Slide Time 8:20 min)



So now the SH is free to linked with any other SH. here we have eight such SH's. So number one can link with number five or it could link with number six or with seven. So we believe that the SH is could practically or possibly linked with any others because there are many such confirmations possible.

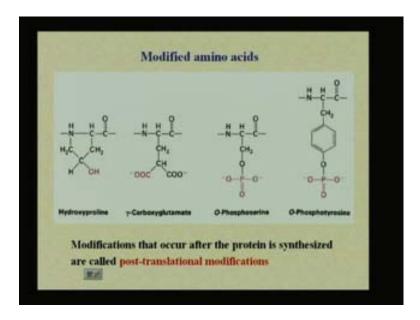
But when we remove the β -mercaptoethanol we are ensuring that these SH linkages or rather the SS linkages can reform. Because we are oxidizing instead reducing it. So if we look at these SH linkages we remove the urea. So we are allowing the formation of SH linkages, we are allowing the protein to refold. Then it is found that it forms exact same structure that was there before. It is exactly just one confirmation and it is the native structure of the enzyme that is fully active with the four disulfide bonds correct to what was there in the original protein.



Now if you just think of the whole thing, it is extremely fascinating which considering that this protein has hundred and twenty four amino acids linked together and has eight SH linkages which could have bonded with any other one which allowed the protein to unfold. So this unfolded structure could have essentially folded into any conformation but it is not happened. It always forms the native structure of the fully active protein with the exact four disulfide bonds correctly linked with one another. It is extremely intriguing about protein folding and it is still not understood to date. And nobody understood how a protein folds into the same structure each and every time (Refer Slide Time 11:27 min). Here we have some modified amino acids. These amino acids are present in proteins to some extent. The names of these amino acids are Hydroxyproline, γ -Carboxyglutamate, o-Phosphoserine o-Phosphotyrosine which are not common amino acids. Here the Hydroxiproline is formed when the Pralines is attached with a specific –OH group which is one other amino acid not among the common amino acids. It can be observed in proteins.

The second one is γ -Carboxyglutamate. Glutamate has only a single carboxylic acid group. The γ -Carboxyglutamate is formed when another carboxylic acid group attached to the Glutamate. And this is a structure which you see in some proteins. Another one in this series is Phosphoserine. In the o-Phosphoserine the phosphate group is linked to the oxygen in the side chain. Here the o-Phosphoserine is a derivative of serine. The serine is now you recognize as part of the polypeptide chain. If you look at this region here is the NH of the o-Phosphoserine in this case and it is the carboxylic acid part of this Phosphoserine moiety.

(Refer Slide Time 13:17 min)



So the linking of these two together by the covalent bond or peptide bond this carboxylic acid group would be linked to the NH of the next amino acid. So here this side chain would be R group. So here we have the Carboxyglutamate is as part of the polypeptide chain. And we notice that we always have an – NH on the left side the amino part of the carboxylic acid always. Because that is the way you read the protein.

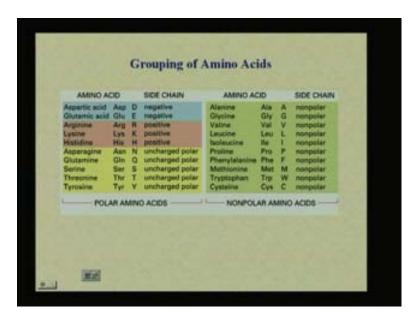
Now you are not always going to get structures you are going to get only letters one after the other. So we have to read the sequence of the protein. We have to recognize the symbols when you read the sequence of the protein. And these are redundant. Here the only information you need to know is about the R group. Because for example if we have Gly- Ala-Val, if we know the structures we can draw the polypeptide chain because we know how they are linked to one another. Here we know the -NH₂ has to be on the Glycine side, the carboxylic part has to be on the Valine side and the Alanine would be intermediate.

So the way we consider all the amino acids is the form of series of letters. So all the information we actually need is a series of letters the amino acid alphabet that comprises proteins or polypeptide chains (Refer Slide Time 15:23 min). Here we have Hydroxyproline in which the hydroxyl group is attached here, we have Carboxyglutamate where the carboxylic acid group is attached here, we have Phosphoserine where the phospho is attached here and we have Phosphotyrosine.

The tyrosine is one of the aromatic amino acid along with the Phenylalanine and Tryptophan which we studied earlier. And so in the Tyrosine we would have phosphate attached to it instead of the OH group. So now it is phosphotyrosine (Refer Slide Time 16:13 min). Usually this is not a common occurrence when the polypeptide chain is formed. All the modified amino acids occur only after the protein is formed. So these modifications are actually occurring after the protein is synthesize and these are called post -translational modifications. Because the process of translation is how you get from RNA to protein.

As we know earlier the central dogma of biology is DNA to RNA to protein. The process of DNA going to RNA is known as transcription, the process of RNA going to Protein is known as translation (Refer Slide Time 17:13 min). So you have the synthesized Protein after the RNA is translating to the Protein. Now if you have the post translational modification you would have the specific modifications that occurred in the already synthesized protein.

(Refer Slide Time 17:42 min)



Now we can see here are grouping of amino acids. These are extremely important in the properties of the Proteins, in the catalysis of the Enzymes, in the catalytic properties of the proteins, in their functional properties, in their surface properties of the proteins. The grouping of these amino acids is very important (Refer Slide Time 18:13 min). Here we will be looking is the glows groupings which means that we will be looking at Aspartic acid and Glutamic acid where these are the acidic amino acids. We will look at Arginine, Lysine and Histidine.

There is one change at times where the Histidine some times occurs in the positive set and the some times occurred in the uncharged polar set (Refer Slide Time 19:03 min). So the ones that definitely negative is Aspartic acid and Glutamic acid. The one's that are definitely positive are Arginine and Lysine. Also we have seen Histidine in this particular grouping of amino acids. What does this mean? If we have a negative amino acid or we have a positive amino acid these amino acids as mentioned earlier are likely to be on the surface of the protein. Because they are going to interact with the solvent where there is a polar solvent (Refer Slide Time 20:07 min). So all the amino acids we are looking here would favorably interact with the polar solvent. The second set that we have here are uncharged polar amino acids. So here the charges come from the specific side chains that they have. Here we can clearly see the changes in the 3-letter codes along with their 1-letter code in the uncharged polar set. Then we have the non polar amino acids. These non polar amino acids would also called Hydrophobic amino acids would tend to be away from the solvent. Because they have no favorable atoms or no favorable interactions with the solvent because all the side chains mostly contain carbon and hydrogen.

So here we have the Alanine, Glycine, Valine, Leucine, Isoleucine, Proline, Phenylalanine, Methionine, Tryptophan and Cysteine. The last three of these have hetero atoms in them. The Methionine has an Sulphur, Tryptophan has nitrogen and Cysteine also has an Sulphur in its side chain. But since they have a predominant amount of carbon and hydrogen to them they are put in a non polar group. So where would these favorably be? Unless they have specific reasons to be on the surface they would rather be bedded in the protein.

Now we have to consider specific properties of amino acids. One thing we have considered is the Size and the shape. Here the Size means molecular mass. The smallest one is Glycine which just has a hydrogen atom and the next smallest one is Alanine because it has the methyl group. Then gradually we can go on to the largest and most the bulkiest one that we could have. Tryptophan is the bulkiest one and Arginine is the longest one. So these are the long chain amino acids that we could have. so the size and the shape of the amino acid is dependant on the R group because the rest of amino acid is same. Here the rest of it we have is the symmetric carbon atom which is referred as the α -carbon atom, we have the amino group, the carboxylic acid group and hydrogen which are common to all amino acids. Then it is only the Size and shape of the R group that differs.

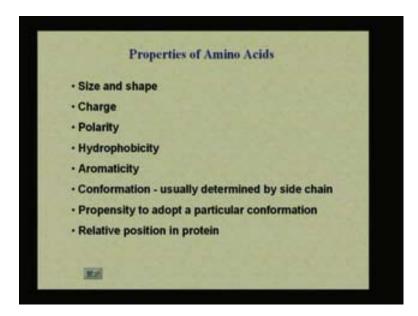
The next one is the Charge. We have already discussed about the groupings. So we can have positive amino acids which are Glycine and Arginine. And we have the negative amino acids Aspartic acid and Glutamic acid. The next property would be Polarity. The difference between the charged ones and the polar ones is that the polar amino acids are preferably some of them being small like serine can occur in the central part of the proteins. They will form hydrogen bonds because they have an -OH to them. Serine has an -OH so it can preferably form a hydrogen bond, Threonine has an -OH so it can also preferably form a hydrogen bond. Here we have two amides and they are Asparagine and Glutamine. So we could have sets of hydrogen bonds formed with them as well. So we could have them being polar forming hydrogen bonds but on charged. Then we have the most important property of amino acids is the Hydrophobicity. It is extremely important because it determines the protein folding nature.

Another property is the Aromaticity. We find three amino acids that are aromatic in nature and they are Phenylalanine, Tryptophan and Tyrosine. And as mentioned previously these three aromatic amino acids contributes to the absorbance of proteins. The optical density or the UV absorbance of proteins is only due to the aromatic amino acid residues. And this we monitored up to 280nm. The UV absorbance of proteins is usually monitored to 280nm. The absorbance or the optical density that is observed is solely due to the presence at 280nm of the aromatic amino acids the Phenylalanine, Tyrosine and Tryptophan. Then we have the conformation. It is usually defined by the side chain. What do we mean by the conformation of the amino acid? When ever we have a long side chain with single bonds in it then it is free to rotate in three dimensional spaces. So we will have a conformation associated with the property of the amino acid.

For example, if we have an Alanine, it just have the methyl group. So that it will not have any conformation because all of will look at identical. But if we have a long side chain like Lysine or Arginine we have a lot of single bonds in them. So these single bonds we can rotate. The NH₂ group that forms the end of the Lysine group will also rotate, would also change as with the single bonds rotation. And as it changes it will also change the conformation of its side chain (Refer Slide Time 27:09 min). This we will study the propensity to adopt a particular conformation. And we will study the relative position in proteins. In which one is the Hydrophobicity. From the Hydrophobicity we can straight away

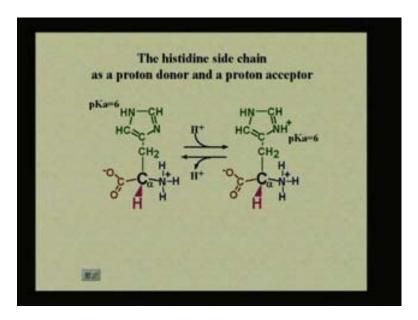
say that for example Phenylalanine or Leucine or Isoleucine is expected to be in the central part of the protein.

(Refer Slide Time 27:45 min)



Lysine and Arginine or an acidic amino acid like Aspartic acid or Glutamic acid is expected to be on the surface of the protein. The Histidine is sometimes considered to be a positively charged amino acid and some times it is not.

(Refer Slide Time 28:28 min)

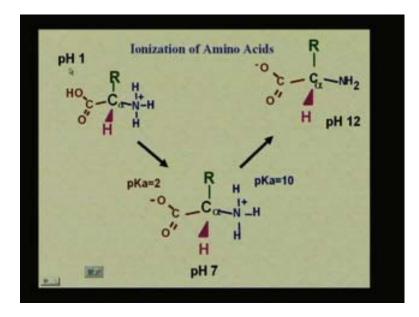


Now, the side chain that you have here is the side chain for the Histidine. Histidine can accept a proton and as it accepts the proton the side chain gets protonated.

So the nomenclature of this Histidine is such that this is asymmetric alpha carbon atom since the carboxylic acid part and the amino part is common to all amino acids. This carboxylic acid carbon is always referred as just C. This C_{α} means that the side chain is always connected to this carbon atom. Then we have this CH_2 group in its side chain in which the carbon atom is referred as β -carbon atom. Here the C_{β} of the side chain is attached to the C_{α} . So in this case the next one would be a C_{γ} . Then we would have two δ -carbon atoms attached to C_{γ} are C_{δ} and N_{δ} and so on and so fourth. So just with the nomenclature alone you can actually identify which amino acid will belongs to because the side chains are different for each case. So an Alanine will have a single C_{β} . Valine will have a C_{β} and two C_{γ} carbon atoms where they are $C_{\gamma 1}$ and $C_{\gamma 2}$. (Refer Slide Time 31:02 min)

So let us move on to the understanding of the proton donor and proton acceptor capabilities of the Hestidine side chain. Here the $p^{Ka}=6$ means when there is an equilibrium at a certain p^H where it can either accept the proton or it can donate the proton. Now, apart from the ionization of amino acids at this particular point the side chain of β -N can protonate to form NH^+ and it can also deprotonate to form this back so this is an equilibrium. This equilibrium occurs at a p^H close or rather between 6 and 7. All the enzymatic reactions in our body are taking place at a physiological pH of 7.4 which where the physiological pH is 7.4

Histidine can form a large number of catalytic sides of proteins is just because of this protonation and deprotonation which means it can accept and donate a proton close to the physiological p^H which is not the case with the other amino acids. If the p^{Ka} is too high then the p^H of the body does not reach that amount of protonation or deprotonation. Because the physiological p^H of our blood or where all the reactions are taking place in our body is close to the p^{Ka} value of the Histidine side chain, this forms a component in many enzymatic reactions in many catalytic sides of proteins. Here we are basically discussing about the ionization of amino acids. So if the amino acid is at a very low p^H then the carboxylic acid group will remain as -COOH and the amine group will remain as NH_3^+ because the solution will have large amount of protons at a very low p^H value. So we have -COOH group and NH_3^+ at a $p^H = 1$ means at a very low p^H value because an acidic solution which has sufficient amount of protons to protonate all the side chains or all the amino groups. But since the amino acid itself has a carboxylic group and amino group so we can consider the protonation of those two sides alone.



Here we have the solution at a $p^H = 1$, we have the protonated carboxylic acid and we have protonated amino group. Here there is no other place where a proton can be accepted. As you increase the p^H or rather adding base by titration the base is going to abstract the protons from the amino acid in the order or in the ease of its abstraction. So the carboxylic acid will loose its proton first. So we no longer have an -COOH group instead of we will have only COO^- at a $p^H = 7$. But the amine group does not loose its proton at this p^H . Keep on increasing the p^H or rather keep on adding more base at a point amine group will loose its the proton which is attached to it. So we will get two inflection points by doing an ordinary titration of the amino acid in which one corresponding to lose of the carboxylic acid H, the other corresponding to the NH_3^+ of hydrogen.

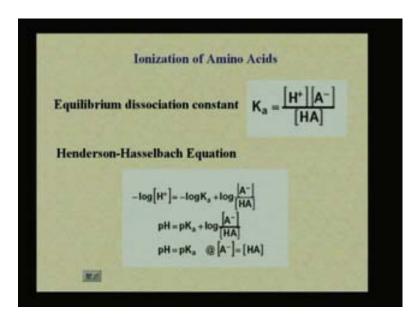
Here the is p^{Ka} is two which means that as the amino acid move close on and reach the p^H close to 2 so the carboxylic acid will loose its proton. Here this p^{Ka} is 10. So the amine group will loose its proton once it reaches a p^H of 10. So at a pH of twelve we have lost all the protons possible in the amino acid. So initially we had everything protonated in the amino acid at a low p^H . As we go on increasing the p^H the amino acid will be in the zwitterionic form. The amine group has a positive charge to it and the carboxylic acid group has negative charge to it so it effectively charged zero. Every amino acid will have this zwitterionic form.

Now what becomes fascinating is, suppose the R group also has a charged so where is this going to be come interesting when I have an acidic group or amine group attached with so this will come into the picture when I have an aspartic acid or glutamic acid or lysine or arginine and the histidine for the pKa six.

We have an equilibrium dissociation constant associated with it if we consider the ionization of the amino acid. This is true for all amino acids. Here the equation we have is HA dissociating into H^+ and A^- and the K_a associated with it is $[H^+]$ $[A^-]$ / [HA]. We will get the Henderson -Hasselbach equation by rearranging the above equation and we write it in such a manner that we get the H^+ on the left hand side. We already know that the negative logarithm of hydrogen ion concentration corresponds to the p^H . So

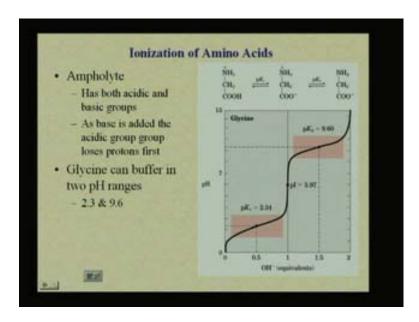
here we have the $p^H = p^{Ka} + \log[A^-]$ / [HA] where this [A $^-$] is the base form, this [HA] is the acid form or other the salt form or the acid form and this is called the Henderson -Hesselbach equation. And if the concentration of these two is equal means we have equal amount of undissociated and dissociated forms. Then this is going to have a ratio of one making the logarithm zero and the pH will be equal to pKa.

(Refer Slide Time 40:52 min)



So here we are considering the amino acid as HA form and we are considering the dissociation from HA to H^+ and A^- . And we will get a p^H is equal to p^{Ka} when $[A^-]$ is equal to [HA]. The titration of an amino acid will have a diagram like this which you will understand very clearly.

(Refer Slide Time 41:45 min)



At first we consider these as ampholytes because they have both an acidic group and a basic group attached to it. This is true for all amino acids you have an acidic group which is the carboxylic acid group and you have a basic group which is the amino group. So these are called ampholytes.

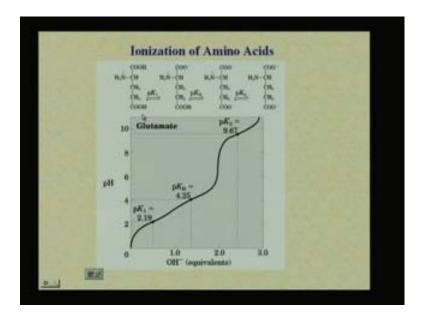
Now as the base is added so this is here along the x-axis we have the number of OH⁻ equivalence added. So we are adding NAOH to the amino acid solution. Initially if we make the p^H of the solution is very low then the amino acid look like all of it is protonated. So we have -COOH and -NH₃⁺. So here I am abstracting H⁺ by adding OH⁻ to it. So there is dissociation immediately as soon as I hit a point this pH. I have equal amount of this species which is my A⁻ and this species.

As soon as we started adding OH^- to the solution this H^+ will starts getting abstracted. So as soon as we started adding the first drop of OH^- does not bring us to the p^{Ka} but it starts abstracting H^+ . So come to a point as we keep on adding OH^- and come to a point where this is equal to this i.e., when the pH is equal to the p^{Ka} . That is on this graph is this point here. So the pK_1 value is 2.34 and at this point we have an equal amount of the carboxylic acid and its anion.

Here this is COO⁻ and this is COOH. Now it is in its zwitterionic form. Keep on adding OH⁻ at some point it will loose the amine proton. The NH₃⁺ has to loose its proton when it cross its p^{K2} then after met this level. So we met p^{K1} where we have equal amount and as we keep on adding OH⁻ we will have more amount of A⁻ than HA. Keep on going with a titration then when we will come to this point of p^{K2} where we do have equal amounts A⁻ and HA. (Refer Slide Time 45:45 min). So this is my H because now this is my acid which has this additional H⁺ added to it. As I add OH⁻ I abstract this proton at a particular p^H the p^{K2} will correspond to a p^H where I have equal amounts of this species and this species. So now if I keep on adding OH⁻ I will just be increasing the p^H because I have just OH⁻ in my solution. So this would be the titration of Glycine. Here you would have two inflection points where the first one corresponding to the carboxylic acid proton and the second one to the amine group.

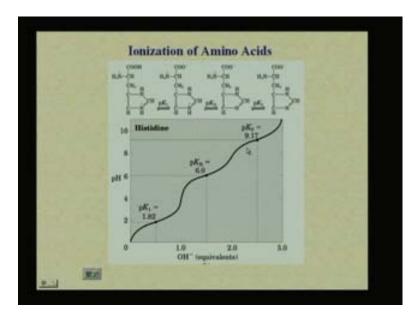
Suppose I have Glutamate. It has an additional COOH to it. Looking at this species this is the side chain. This is the carboxylic acid part of the amino acid, this is NH_3^+ . So at this point all of them are protonated because the p^H is very low. So all of these are protonated at this region where it has a very low p^H value. Now I am adding OH^- . Then what will happen? I will start abstracting the proton or rather the group of that will most easily lose its proton will start the equilibration. So that will be this proton, the proton that will belongs to the carboxylic acid of the amino acid not of the side chain which will be lost first.

(Refer Slide Time 48:29 min)



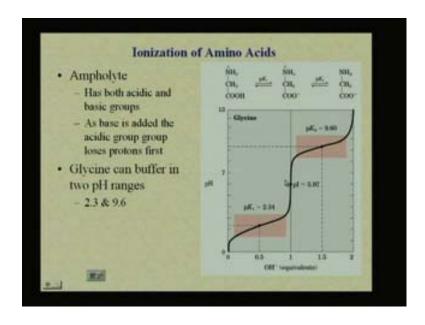
So what is the species that is being formed? It is this. This is still protonated and also this is still protonated. Now, I am continuing with the titration. How many more protons that I could loose? Two, one belonging to the COOH of the side chain and one is the NH_3^+ proton. So as I go on further with OH^- addition in the titration. The next easiest one to be removed is the side chain proton. Now I keep on titrating then what will happen is I have to loose the last proton which is left in the amino group. So the different pK values where these have correspond to the three protons that can be lost. We have three protons that could be lost. The first proton which we have lost is the carboxylic acid group attached to the C_{α} of the amino acid but not the side chain because that is the most acidic proton. Then we move to the next one, which is the next most acidic proton? The side chain one. So we have the side chain proton that we are now going to loose. Then we loose the amine group one. So what do I have in this curve? I have a pK₁ value, I have a pK_R value where R is the side chain so it corresponds to the side chain. Then we have a pK₂ value which corresponds to the amine.

(Refer Slide Time 51:20 min)



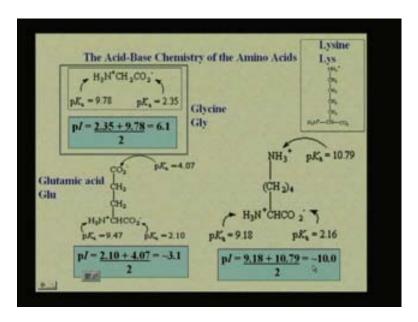
Now let us consider Histidine. Now we have all the possible sites are protonated at low p^H . So we have protonated side chain, we have protonated amine group of the amino acid and we have protonated COOH group. We will loose this proton as we increase the OH⁻ additions. So as we loose this proton, the next proton that we are going to lose is the one belonging to the side chain. The last one that we are going to loose is the one that belonging to the most basic group which is the amine group. So again I have the titration curve rather look like. So I have a pK_1 corresponding to the COOH, I have pK_2 corresponding to the amine group. Here we have two pK_1 values for the Glycine. What is this form? It is the zwitterionic form. So at a particular point sometimes we have a pK_1 value and some times we have pI value.

(Refer Slide Time 53:27 min)



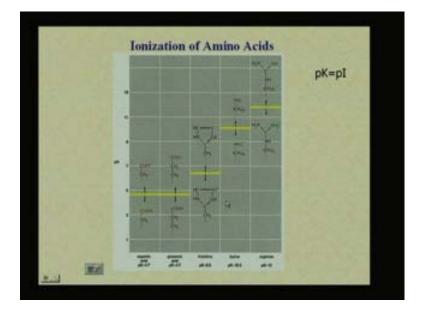
That is the Isoelectric point. The Isoelectric point is the point at which you would have the zwitterionic form of this amino acid and the charge would be zero. So what is the pI value? It is the sum of the pK_a of carboxylic acid group and the pK_a of the amine group divided by 2. What happens if I have a Glutamate? The charge is zero for the Glutamic acid when it is between these two. Now see this is rather pK_2 for the amine, this is pK_R and this is pK_R so it first loses this one.

(Refer Slide Time 54:42 min)



So the charge is 0 in between these two. So the pI value is [2.1 + 4.07] / 2 = 3.1. It depends on what p^H of your solution it is. If you have a p^H corresponding to the pI then the charge on the amino acid is 0. Here we have Lysine which has an additional NH_3^+ , when you study a titration curve of Lysine it is going to be have differently and we have the pI at a point where Lysine will be uncharged that is between the pK_a of the amino group and the pK_a of the side chain this time. So the pI of the Lysine is very high.

(Refer Slide Time 55:55 min)



In summary, we have the ionizations of the amino acids, these are regions called the low p^H regions corresponds to the Aspartic acid, Glutamic acid and the intermediate region to the Histidine then we have lysine and finally we have Arginine. So these Lysine and Arginine are the two basic amino acids, Aspartic acid and Glutamic acid are two acidic amino acids and this is as I said act as a proton acceptor and a proton donor depending on the value of p^H. So we will stop here for today. So the discussion that we have today was on understanding the properties of the amino acids their size, their charge, their polarity, and the titration of the amino acids which is extremely important when we consider the ionization that could occur based on the pH of the solution.

Thank you