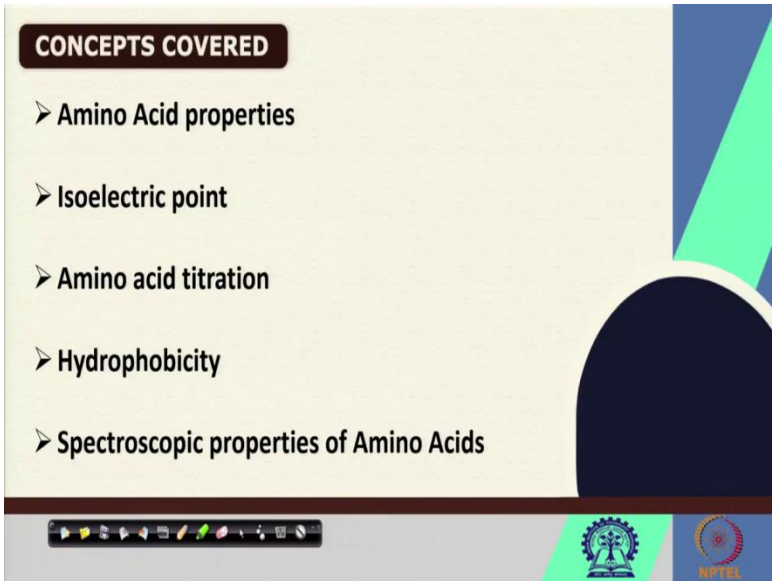
Fundamentals of Protein Chemistry
Prof. Swagata Dasgupta
Department of Chemistry
Indian Institute of Technology, Kharagpur

Module – 01
Amino Acids and the Peptide Bond
Lecture – 02
Amino Acids – II

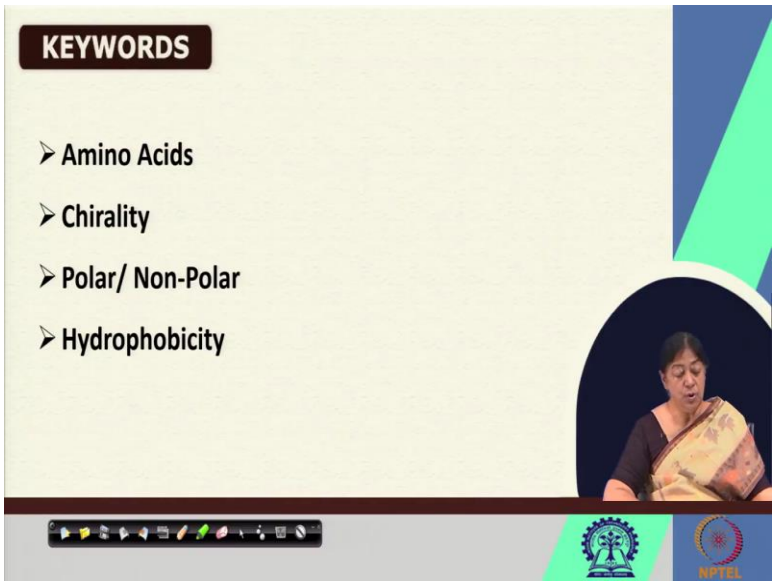
In the 2nd lecture of module 1 we will talk about Amino Acids and their properties and further to the types of amino acids that we learnt in the last lecture, we will see what further characteristics are there of these unique building blocks of proteins. To understand these we have listed a set of properties.

(Refer Slide Time: 00:38)



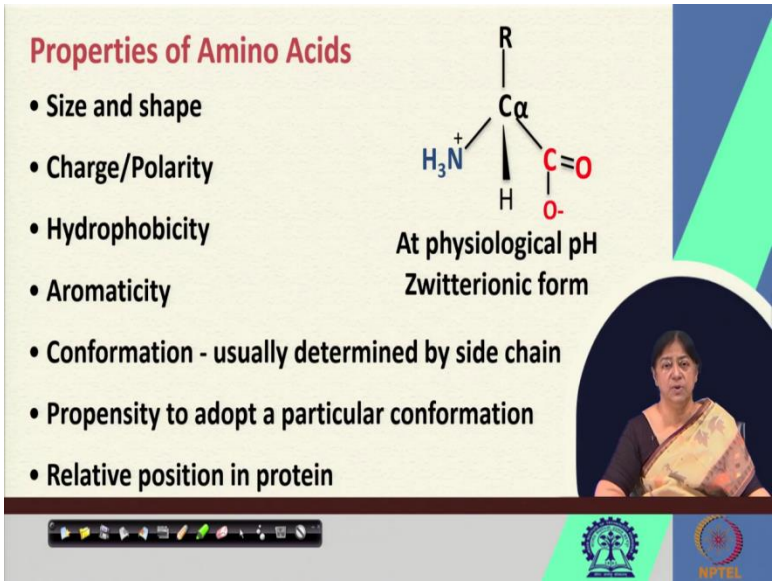
The isoelectric point, an amino acid titration, which is going to help in an understanding of how we look at amino acids, what their structural properties are, the hydrophobicity of the properties of the amino acids and the spectroscopic properties, which we will do in the next lecture.

(Refer Slide Time: 00:57)



Now, to look at amino acids we understood what a chirality meant; the polar and the non-polar parts and most important the hydrophobic characteristics of these amino acids.

(Refer Slide Time: 01:09)

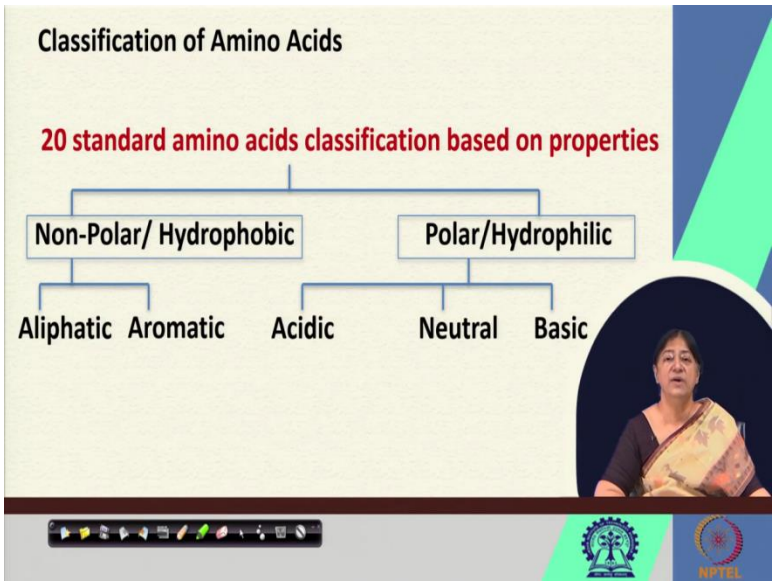


In the general properties of amino acids, when we look at what we call the zwitterionic form, we know that this is a neutral amino acid. Why, because it has an NH_3^+ and a COO^-. Now, this zwitterionic form is the most common form that is going to be present at the physiological pH of 7.4.

We look at the size and the shape, because these are going to be important in how it interacts or how it appears in a protein. We are going to look at the charge, the polarity, the hydrophobicity, the aromaticity, the conformation that is usually determined by the side chain and the propensity to adopt a particular conformation.

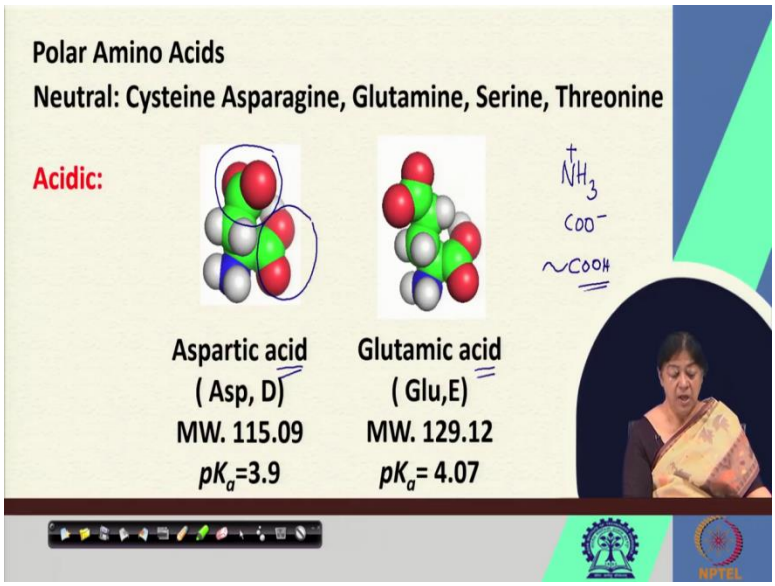
Now, as we go on in these different properties and see how they affect proteins in general, not only in structural characteristics, but definitely in functional characteristics as well. And of course, their relative position in the proteins is going to depend upon the type of protein that we also look at.

(Refer Slide Time: 02:19)



The 20 standard amino acids that we looked at in our class last time, was that we have the non polar or the hydrophobic type and we have the polar or the hydrophilic type; where we have an aliphatic, aromatic set or we have an acidic, neutral and basic set.

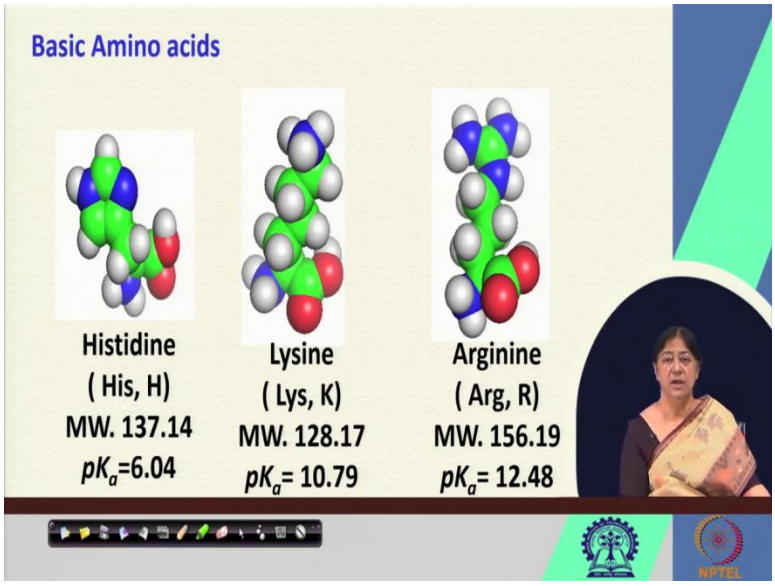
(Refer Slide Time: 02:36)



If we go into further details and try to understand what these structures are all about, then we see that in the aspartic acid and the glutamic acid, that are the acidic amino acids, there are larger numbers of red atoms, meaning there are larger number of oxygen atoms.

Simply because these are the acidic amino acids, which means that apart from the NH_3^+ group and the COO^- group that every amino acid has, this has in addition, a side chain that is also an acidic amino acid. Now, when we look at this particular set, we have to try and understand that this plays an important part.

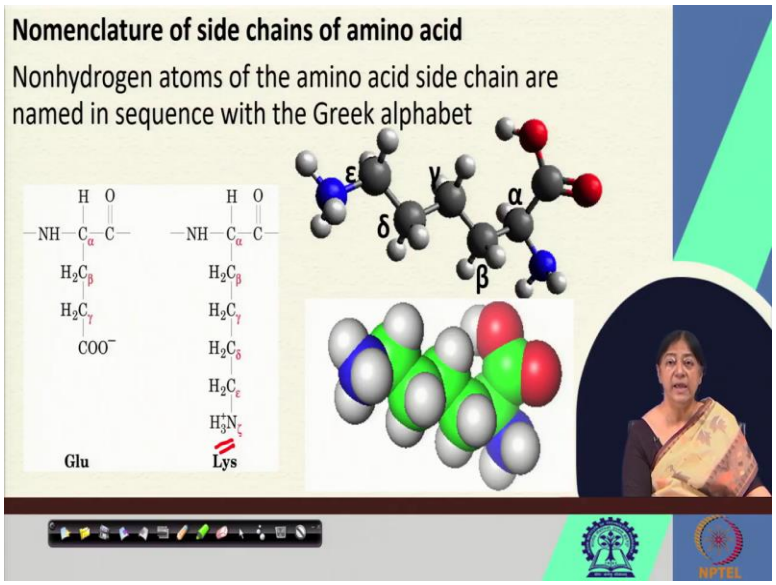
(Refer Slide Time: 03:33)



We look at the basic amino acid residues and we see that when we look at these residues, these are going to help us in trying to understand what we mean by the side chain residues. What are these side chain residues? The side chain residues are going to be those that are going to give us an additional nitrogen in them.

So, whether it is histidine that has an imidazole group, whether it is lysine that has the epsilon NH_3^+ group or the arginine that has a guanidine group. So we see the extra nitrogen atoms here. Now, how does this help us in understanding? We will see as we go along, not only in the interactions but also in very important aspects of proteins.

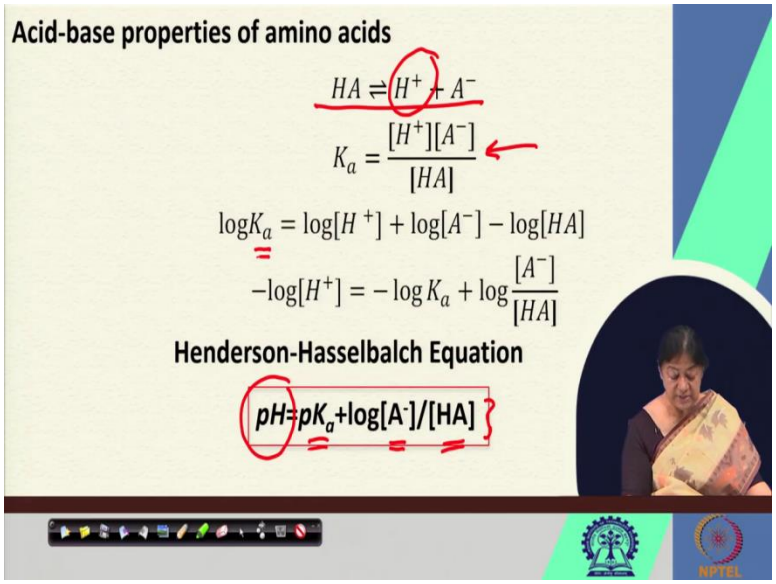
(Refer Slide Time: 04:20)



So, when we look at the amino acids, we talk about the α carbon, but there are many other chains or many other atoms, moieties that are attached to this central α carbon and we have to look at them in terms of a nomenclature. In trying to understand in a universal fashion, how we depict these amino acids in general.

So, if we look at the lysine that has been drawn here, we can see that there is a Cα carbon atom, then there is the Cβ, the Cγ, the Cδ, the Cε and the NH_3^+, which is the side chain of the amino acid. So, when we look at the nomenclature or we look at the way they actually form, this is what is important in an understanding of the amino acid characteristics or the amino acid depiction and the way we name them.

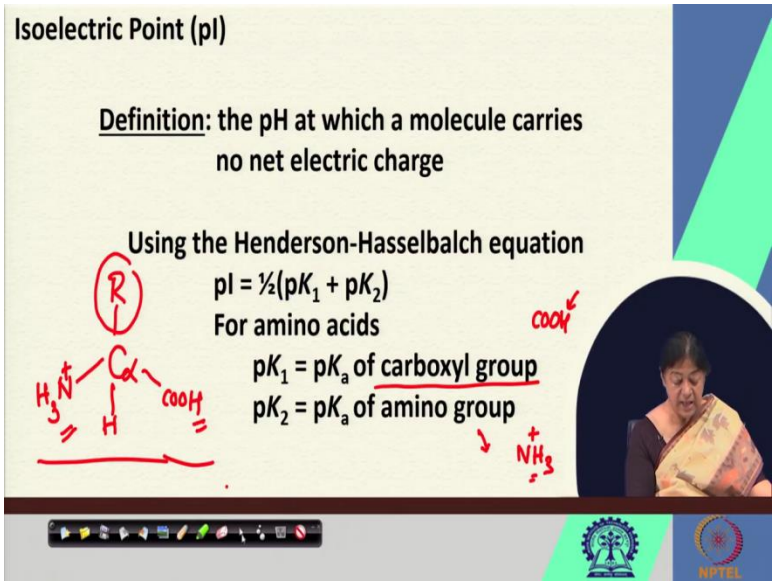
(Refer Slide Time: 05:17)



When we look at now, the acidic and basic properties of amino acids; we have the NH_2 group we have the COOH group and we know in the zwitterionic form this is going to remain as NH_3^+ and COO^-. What does this mean? This means that, these protons that are present can be lost.

The loss of a proton means there is a particular reaction that can occur that is going to tell us or take us from HA to H^+ and A^-, where we have a conjugate base and we have the loss of the proton. Now, this loss of the proton, using the law of mass action in an equilibrium setting, we can find out what the value of the K_a is, and from the K_a value we get what is called a pK_a value and from that, with a concentration knowledge of the A^-, that is the conjugate base and the HA, we can find out at a definite pH what is this ratio going to be. Now, how is this going to help us? What can we learn from this understanding? So, when we look at an idea of this particular setting where we are losing a proton, we are trying to look at how the proton is affected, how the amino acid is losing this proton.

(Refer Slide Time: 06:53)



We will look at the definition of isoelectric point. It is the pH at which the molecule carries no net electric charge. We will see how this is important later on when we study more about proteins and their overall distribution of charges, their overall characteristics and their overall interactions.

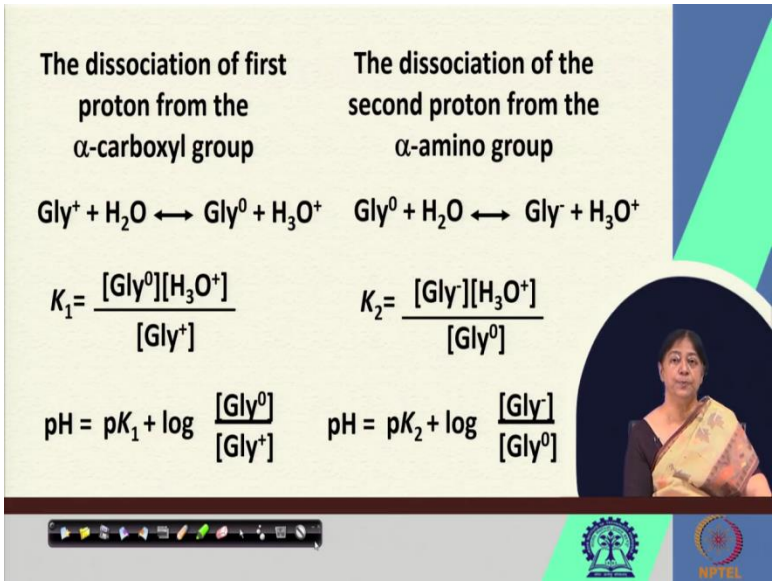
So, if we use this Henderson-Hasselbalch equation, we know that each amino acid has an amino group and a carboxyl group. And what we are trying to do is, to find out the isoelectric point that is actually the average of the pK_1 value, that is the pK_a of the carboxylic group, where in the COOH, a proton is lost and the pK_a of the amino group where we have the NH_3^+, where we know a proton is lost.

As we know, every amino acid has a Cα carbon atom, a H, a NH_3^+ group and a COOH group. So this would be the characteristic of an amino acid at high acidic condition, which would mean that

all of the possible places or all the possible atoms are protonated. And this R group will then add additional acidity or basicity to the amino acid, which is what is important.

So, if we try and understand this further, then we will have to see what we mean by these basic characteristics. So, let us look at the next set and try and understand what this amino acid characteristic is.

(Refer Slide Time: 08:57)



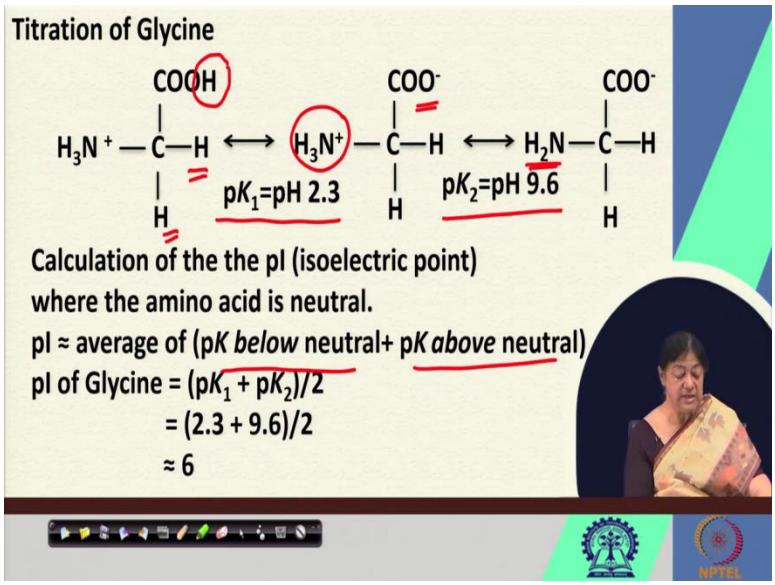
So, we have the dissociation of the first proton from the α-carboxyl group. This is going to be followed by the dissociation of the second proton from the amino group and this is common for all amino acids. So, if we look at the simplest amino acid; glycine, it has a plus charge. Where is this plus charge from? It is from the NH_3^+, because the COOH that is at the lower pH does not have a charge to it yet.

Now, as we have the loss of the first proton, the COOH is going to lose its proton and become COO^- in forming the zwitterionic form. So, we have a K_1 value associated with the loss of this first proton from the carboxyl group which is associated with a pK_1 and a pH.

Similarly, if we lose the second proton from the amino group this time, which is NH_3^+ going to NH_2. We will have a similar expression relating to the K_2 and we will get:

$$pH = pK_2 + \log \frac{[Gly^-]}{[Gly^0]}$$

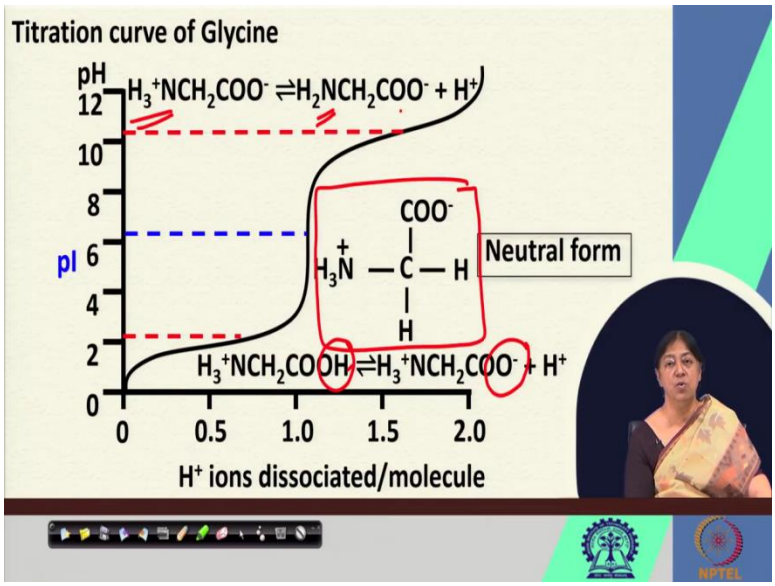
(Refer Slide Time: 10:11)



Now, when we look at the structure of glycine. We know that it has two hydrogen atoms to it, the simplest amino acid. We have the pK_1 and pK_2 values. Now, when we look at these values here, the first loss is going to be the proton in COOH, the second loss is going to be the proton in NH_3^+, resulting in COO^- and NH_2. So, what does this tell us?

This tells us that the pI is going to be the average of the pK below neutral to the pK above neutral. And what is neutral? 0. So, we look at pK_1 plus pK_2 divided by 2, that is going to give us the pI of glycine, fine. So, now we are looking for a pK value below neutral a pK value above neutral and that is going to tell us the pI of the protein.

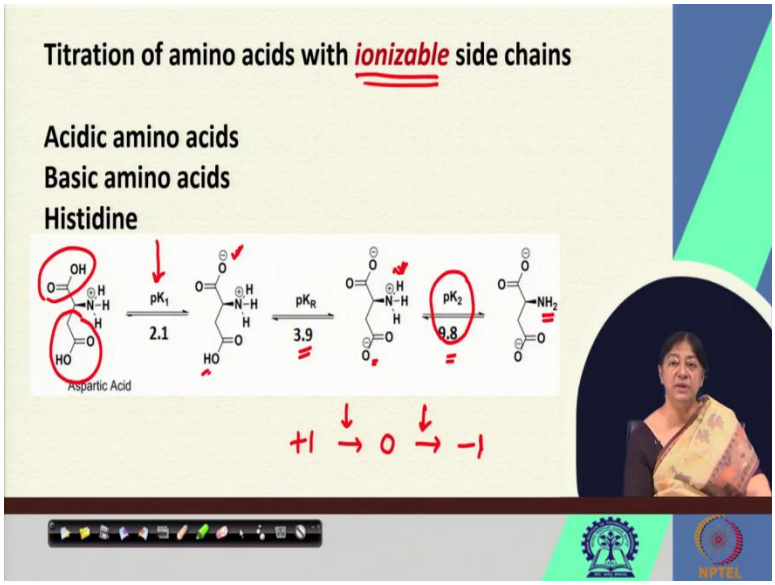
(Refer Slide Time: 11:05)



So now, if we look at the titration curve of glycine, we have the first break where we have the H loss to COO^-, forming what is called the neutral or the zwitterionic form here. This then goes on

to lose its next proton. So NH_3^+ going to NH_2 giving us the final form of this species in the titration, where we are losing the H^+ ions as we go along by adding say sodium hydroxide.

(Refer Slide Time: 11:43)

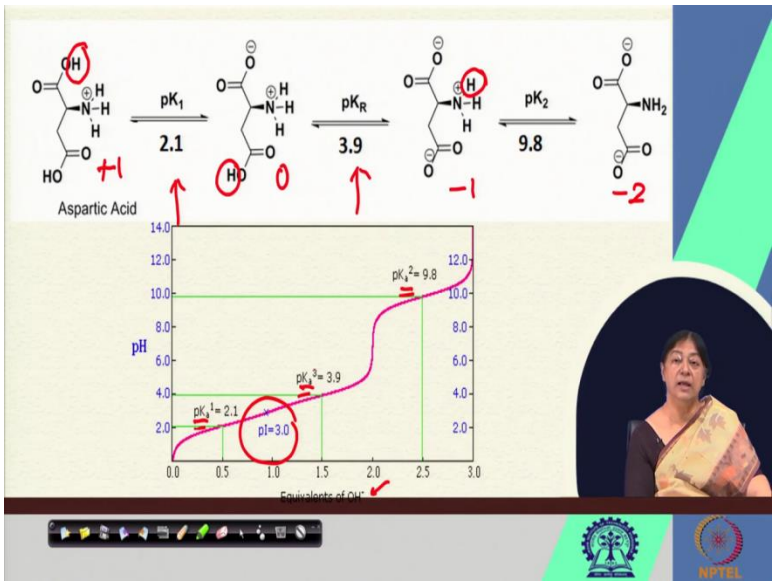


So, when we look at the titration of amino acids now with ionizable side chains, this is going to make it a bit different. Why? Because the acidic amino acids are going to have in addition to the general NH_3^+ and the COOH that is associated with every amino acid, they are going to have an additional side chain. So this side chain of the aspartic acid that we are looking at also has a proton, which will be lost when we do the titration.

So, the first proton loss depending on the acidity or how easily the proton is lost, we are going to first lose the proton in the COOH group. So, at the first instance pK_1 we lose this proton. At the second instance which is the pK of the R, meaning the pK of the side chain, we are going to lose the proton of the COOH group attached to the side chain.

Now, we see we have lost this proton. And at the end at the pK_2 value, this corresponds to the NH_3^+ which is going to be NH_2. So, what does this mean? This means that we have to go from below neutral, say from a charge of +1 to 0 to -1 and the pK_a values that are associated with these changes, the average of those is going to give me the pI of the particular amino acid.

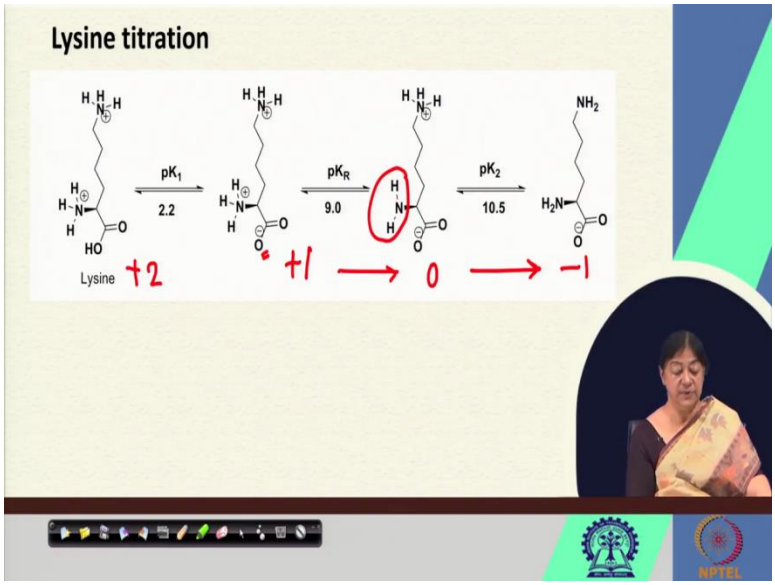
(Refer Slide Time: 13:17)



So this is the way we look at it. We have aspartic acid, we have a pK_1 value. Now that we understand this, we have a loss of the proton in the COOH group of the amino acid first, then a loss of the proton of the COOH group attached to the side chain and finally a loss of the proton in the NH_3^+ group. So we know that the overall charge here is +1, the charge here is 0, the charge here is -1, the charge here is -2.

And what did we learn? We need to get the average that is going to take us from +1 to 0 to -1. So we need the pK_1 value and the pK_R value to tell us what the pI of aspartic acid is. So, this is a simplistic way of looking at the titrations. So we are adding equivalence of OH^-. We are looking at a titration where we have pK_a^1=2.1, pK_a^3=3.9 and pK_a^2=9.8 [refer to slide] and we understand what this is for. So now we have a pI that is going to be the average of the pK_1 and the pK_R.

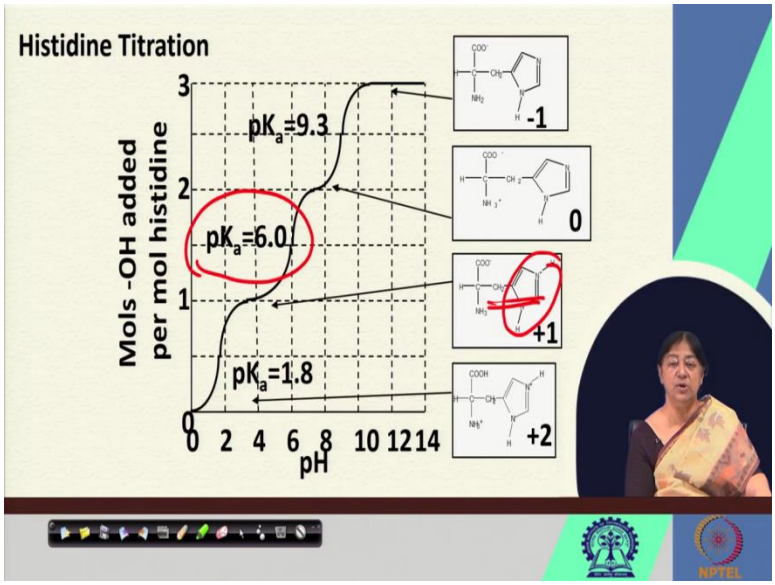
(Refer Slide Time: 14:22)



Similarly, if we look at lysine titration, [Refer to slide] at first we have a charge of +2, and then upon losing a proton in the COOH group we get a charge of +1. Now, when we look at the second set, when we are going for the pK_R, the loss of a proton of the NH_3^+ group, we get a charge of 0.

Now, when we go to the next case we have a charge of -1. So, we see the charge change from+1 to 0 to -1. So, now we have to look at the average of the pK_R and the pK_2 in case of this basic amino acid, lysine.

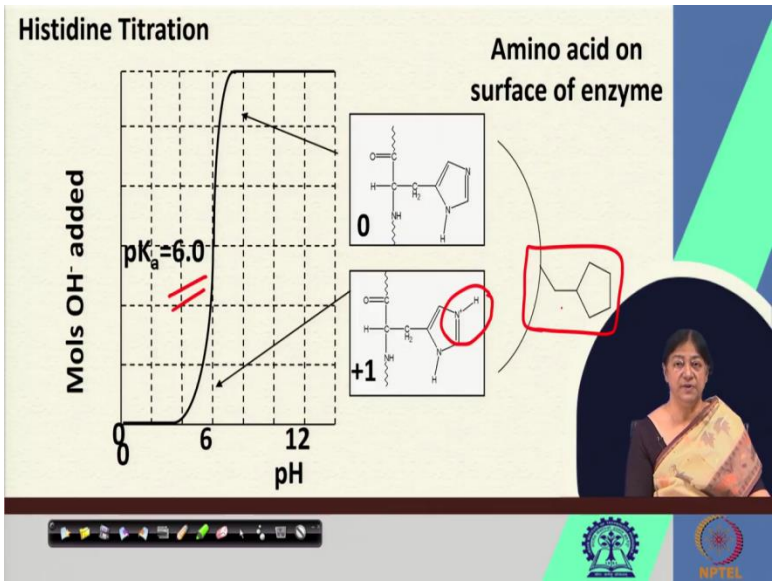
(Refer Slide Time: 15:08)



Similarly, we can look at other amino acids as well. The interesting one is histidine. The histidine titration is interesting simply because if we look at the pK_a value it is 6. Now, why is it interesting? Because, the pK_a value is going to change or it can change depending upon how easily the proton is going to be lost from the side chain. In this case, it is the imidazole group.

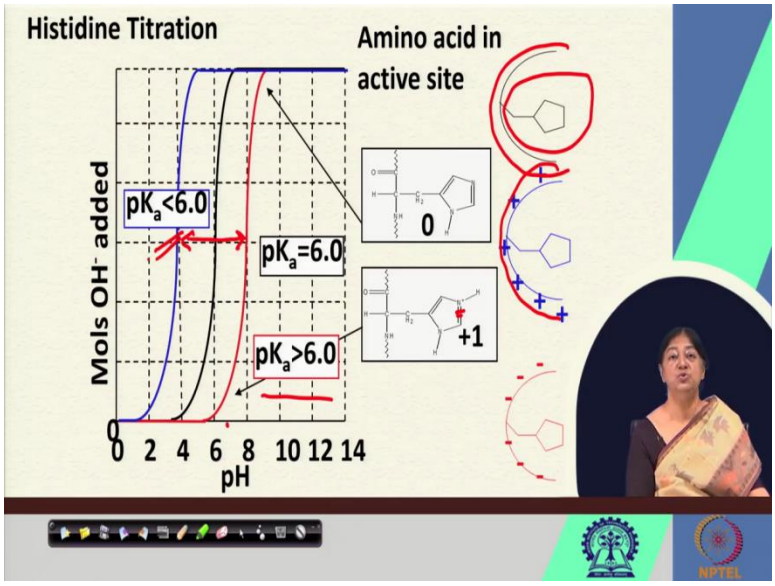
When we are looking at the imidazole group that is connected to Cα, which is the amino acid histidine, we see that it is the most interesting one. Now, why is this important? This is important because the pK_a value of this particular side chain of the imidazole group has a pK_a value of 6. Now, the pK_R of the side chain depends upon its environment. And, what we will see later on we when we go into the depths of understanding of proteins, we will see that histidine occupies the active sites of many enzymes.

(Refer Slide Time: 16:19)



So, it is important to see the role of this histidine in terms of a titration. So, if we consider the imidazole of the histidine at the surface of an enzyme, we will see that the pK_a value is 6. Now, what does this pK_a value correspond to? It corresponds to the loss of this proton of the imidazole group.

(Refer Slide Time: 16:46)

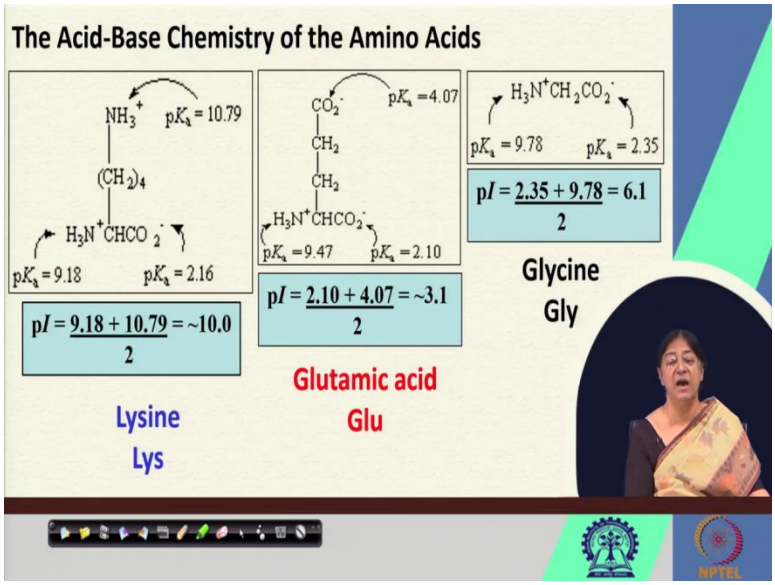


Now, if we look at a rough representation of the imidazole, looking at it on the surface of an enzyme. Now, what happens if this amino acid is at the active site? So, here we have a rough sketch of a protein and here is the imidazole group in the active site of a protein. So, it can lose its proton, but will the loss depend upon what is surrounding it? Yes, it will. Suppose we say that there is an abundance of positive amino acids close to this histidine, it will not want to hold on to the proton, because that renders it a positive charge.

So it would rather release the proton at an earlier point in the pH and then what would happen is the pK_a value would be less than 6. In the same way, if we say that this is surrounded by negative amino acid residues, in abundance we have acidic amino acid residues, it would rather hold on to the proton and prevent its loss as much as possible rendering a pK_R value that is greater than 6.

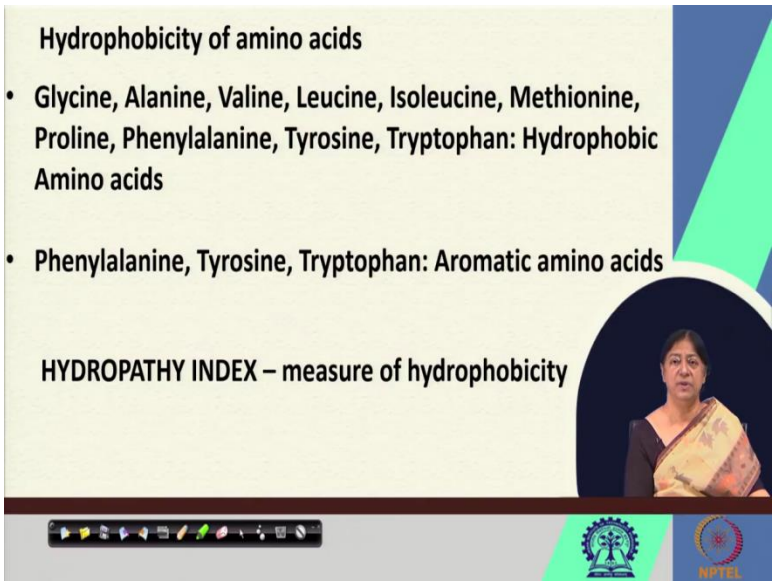
So, when we have a greater than 6 value, we can have a less than 6 value and the interesting part here is, the range of this covers our physiological pH of 7.4. It is interesting to note that the variation in the histidine can go from 5 to even close to 8. So, that is an important feature which we will understand later as we go along.

(Refer Slide Time: 18:26)



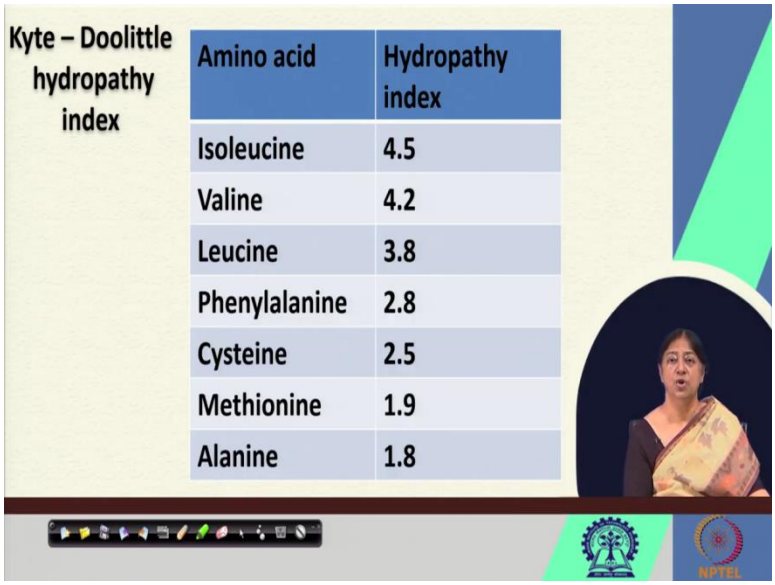
So, if you look at the acid base chemistry of the amino acids, we saw how glycine behaves, we saw how an acidic amino acid behaves and we saw how an basic amino acid behaves. So we know now how to calculate the pI of an amino acid and later on when we study the peptide bond, we will understand how we can actually look at a peptide bond or calculate the pI for a series of amino acids.

(Refer Slide Time: 18:55)



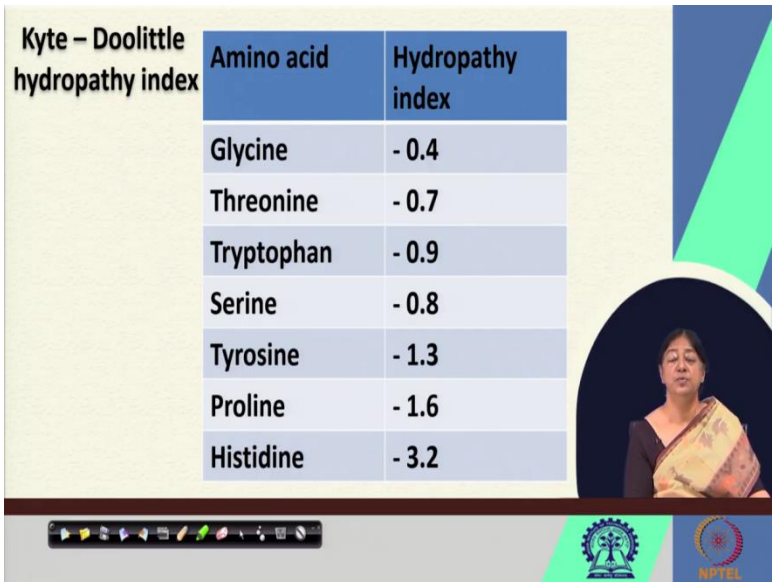
The next very important property is the hydrophobicity of amino acids. Now, this is extremely important in understanding where an amino acid is going to be comfortable in a folded protein, which we will understand as we go along. But the property in trying to understand what is present in the side chain of these amino acids, we learn that the hydrophobic amino acids preferably have only carbon and hydrogen to them. The measure of hydrophobicity is given by a hydropathy index.

(Refer Slide Time: 19:34)



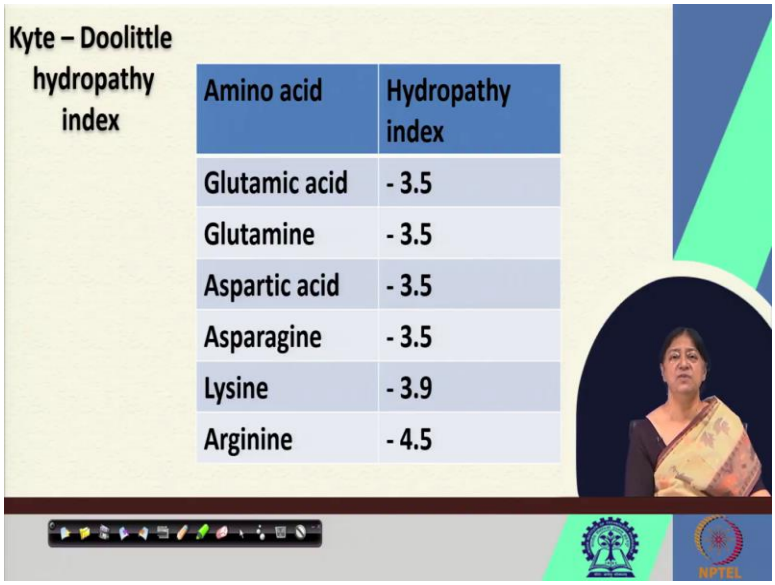
There is a common index known as the Kyte - Doolittle hydropathy index, which gives us as you can see, positive values for highly hydrophobic amino acids; isoleucine, valine, leucine, which are the hydrophobic amino acids meaning their side chain contains only carbon and hydrogen with no heteroatom, that is no nitrogen, no oxygen and no sulphur.

(Refer Slide Time: 20:00)



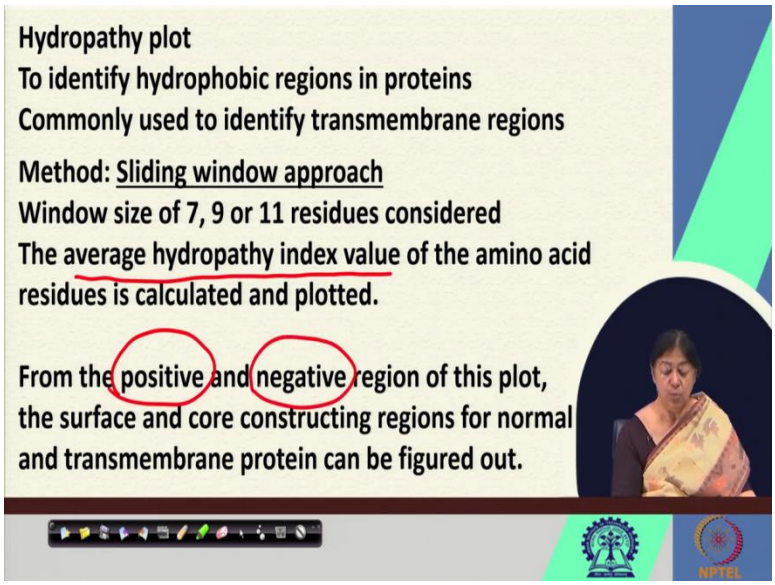
As we go on to the other types, we will see the values gradually decreasing saying that their hydrophobicity decreases.

(Refer Slide Time: 20:09)



And as is expected, with the acidic and the basic amino acids we are going to have high negative values. Now, how does this help us?

(Refer Slide Time: 20:24)

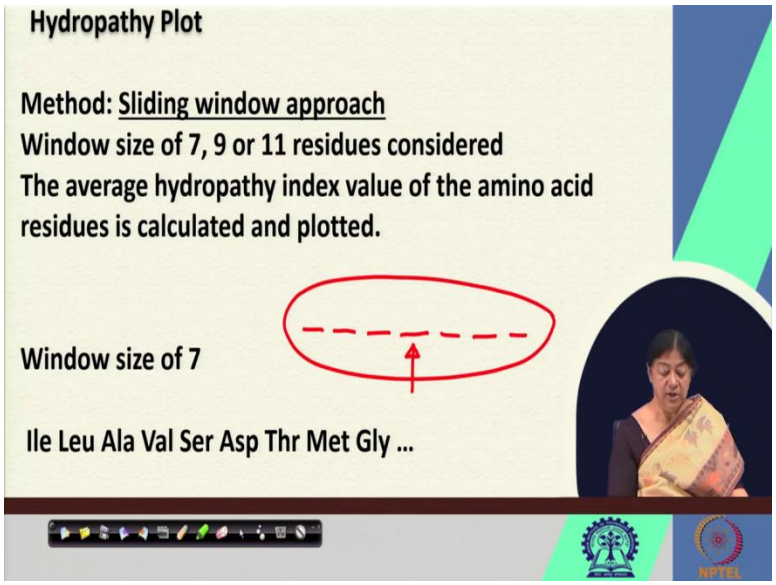


This helps us in a way that, if we have amino acids one after the other in a particular series, we can actually identify hydrophobic regions in proteins and this is commonly used to identify transmembrane regions. We will look at membrane proteins as well, following transmembrane regions.

So, we have what is called a method called the sliding window approach, which we will just address in a simple manner. Looking at a specific window size we will explain what this means in a moment. Calculating what is called an average hydropathy index value of the amino acid residues and we plot this.

Now, from the positive and the negative region of the plot we can determine the surface and the core regions for proteins. So, we know when we saw the table that a positive region means we have a hydrophobic region and a negative region would mean we have a hydrophilic region.

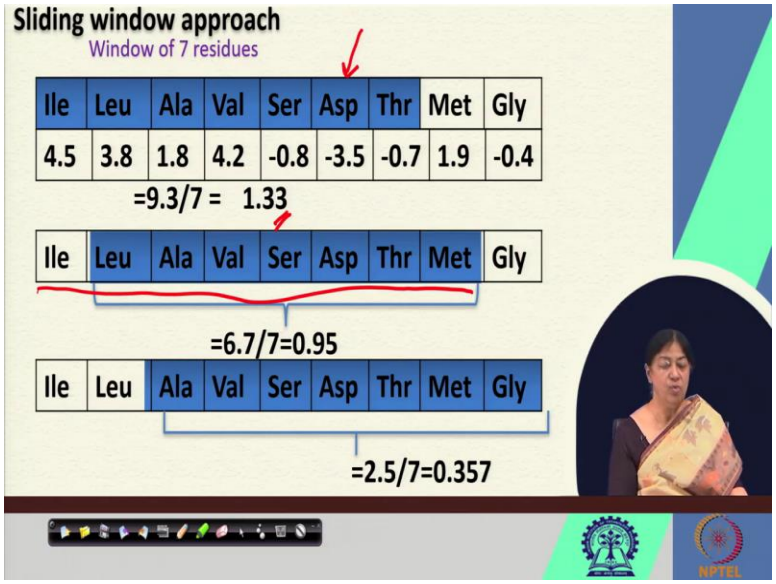
(Refer Slide Time: 21:20)



So, let us look at what this looks like. We have the sliding window approach, 7 or 9 residues and we look at the hydropathy index; we look at a window size of 7. Now, window size of 7 means we have 1, 2, 3, 4, 5, 6, 7. We calculate the average of the hydropathy index values from the table that we showed and we assign it to the middle residue here. Then we slide the window as shown below.

So, say we have a sequence like, what is given here what are these isoleucine, leucine, alanine, valine, serine, aspartic acid, threonine, methionine, glycine. So there is a mixture of acidic, neutral, but polar and hydrophobic residues. Let us see what the characteristic of this string looks like.

(Refer Slide Time: 22:17)



So, here we have the string and the particular hydropathy index values have just been assigned to them from the table. And what we do is, we look at the first seven in this case and we add up the values, calculate the average, because we know our window size is 7 and we say that the value is 1.33 positive value.

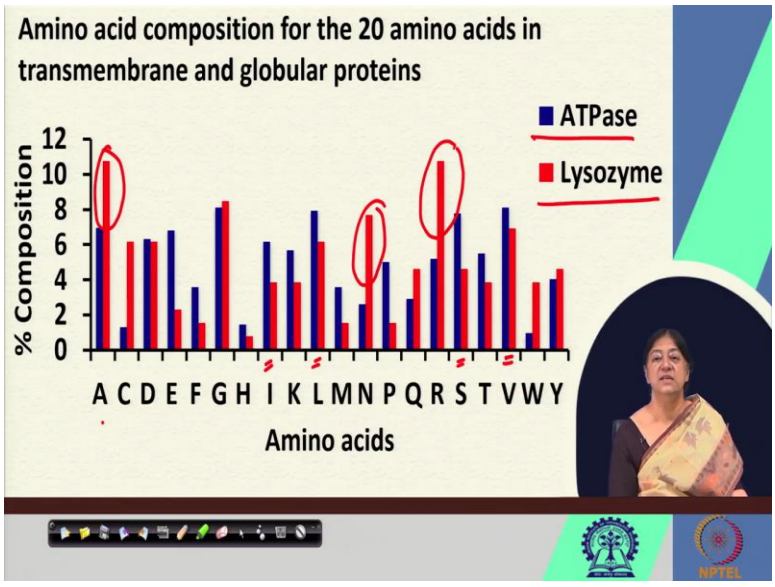
Then we look at the sequence again, what we do is, now we look at our specific amino acid. But now we have to slide over one, because it is a sliding window approach, so we still have 7 residues in our window.

What do we do? We again calculate the number, the total number, then divide by 7 to get the average and we still see that the value is 0.95. And then we go on to do this for the next set and we keep on doing this till we cover the whole protein right. Now, when we do this, what we are actually trying to look at, is whether we have a positive region.

Now, what is the positive region going to tell us? The positive region is going to tell us that we have a hydrophobic region in this case because, if we notice here, we see an acidic amino acid, but the overall setup or the overall region that we are looking at actually is hydrophobic in nature, despite the fact that we have a polar amino acid and an acidic amino acid.

Now, the window size of 7 or 9 or 11 is sometimes chosen, because we can assign the value to the middle residue. We do not want too small a window size because that is going to create noise in our plot and we do not want it to be too large because that we might over compensate for the values and not be able to get an idea of what this particularly looks like.

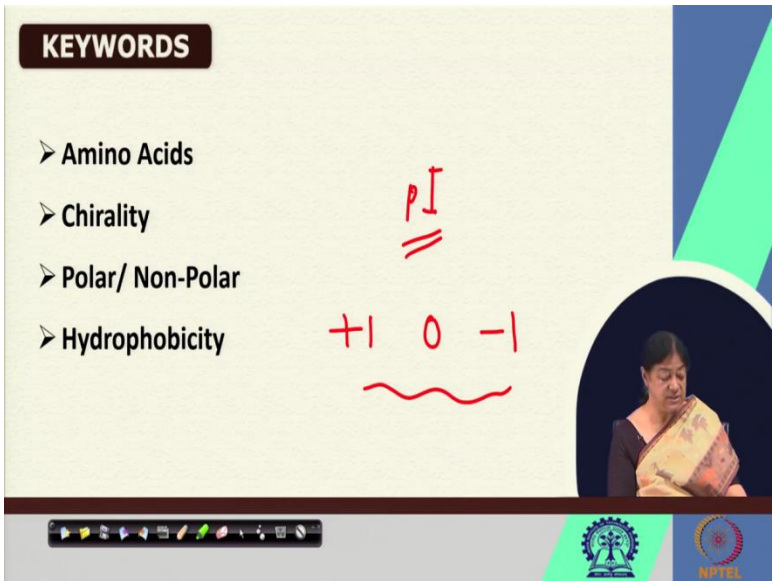
(Refer Slide Time: 24:25)



So, here we have the amino acid composition for 20 amino acids in transmembrane and globular proteins. So what does this look like? This looks like two proteins ATPase which is a membrane protein and Lysozyme, which is regular globular protein. So, if we try to look at the different types of amino acids that we have here and where they are located.

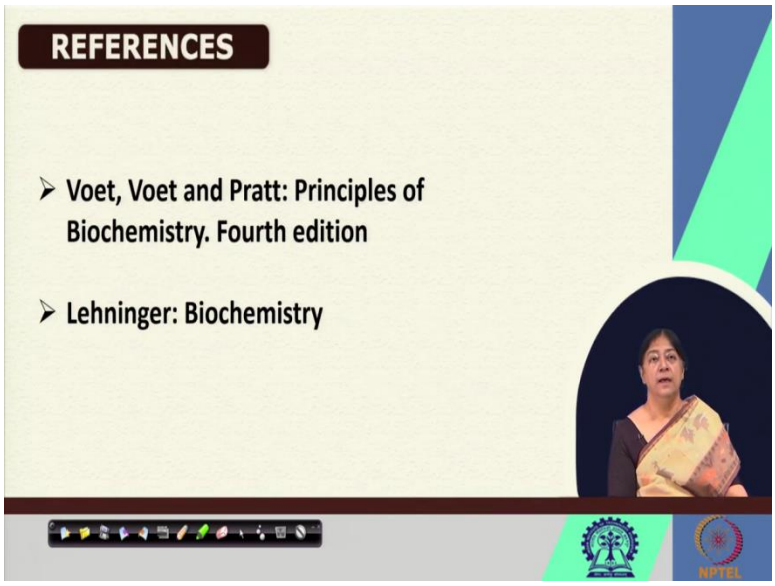
Let us look at the larger ones here: arginine, alanine and asparagines [Refer to slide]. But, if we look at the blue ones where we have valine, serine, leucine and isoleucine, these are the ones that are hydrophobic in nature. So, the hydrophobic nature tells us that it is preferably a transmembrane protein.

(Refer Slide Time: 25:21)



So, what we learned here was the amino acids, their chirality, the polar, non polar aspects of them and the hydrophobicity and an understanding of how we can calculate the pI, the isoelectric point of a protein. What is this isoelectric point? It will tell us taking us from a charge of +1 to 0 to -1, telling us that this is important. And, when we looked at histidine, we found out the importance of histidine in terms of its environment as to how easily it is going to lose the proton that is associated with the imidazole group of the side chain that is histidine.

(Refer Slide Time: 26:10)



These are the reference books.

Thank you.