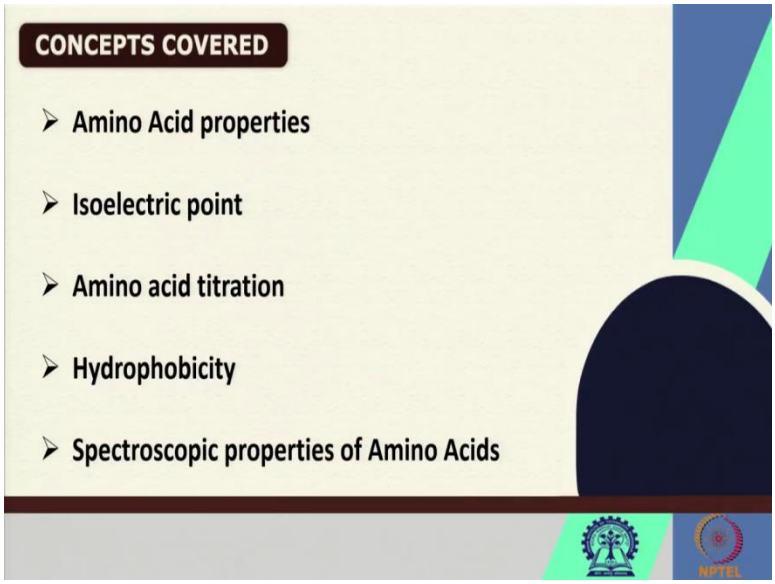
Fundamentals of Protein Chemistry
Prof. Swagata Dasgupta
Department of Chemistry
Indian Institute of Technology, Kharagpur

Module – 01
Amino Acids and the Peptide Bond
Lecture – 03
Amino Acids – III

(Refer Slide Time: 00:23)



We will continue our discussion on amino acid properties and in this particular class we are going to understand the spectroscopic characters. In the ones that we considered previously, we looked at amino acid properties in general, in terms of the side chains, in terms of the presence of the heteroatoms; whether we had only carbon and hydrogen for the side chains or whether we had oxygen and nitrogen or sulfur in them.

So, we need to know which amino acids are ionizable, which amino acids have basic properties or acidic properties. And in the last class, we looked at the isoelectric point in amino acid

titrations and we tried to understand where this neutral part of an amino acid comes from in terms of what we know as the zwitterionic form.

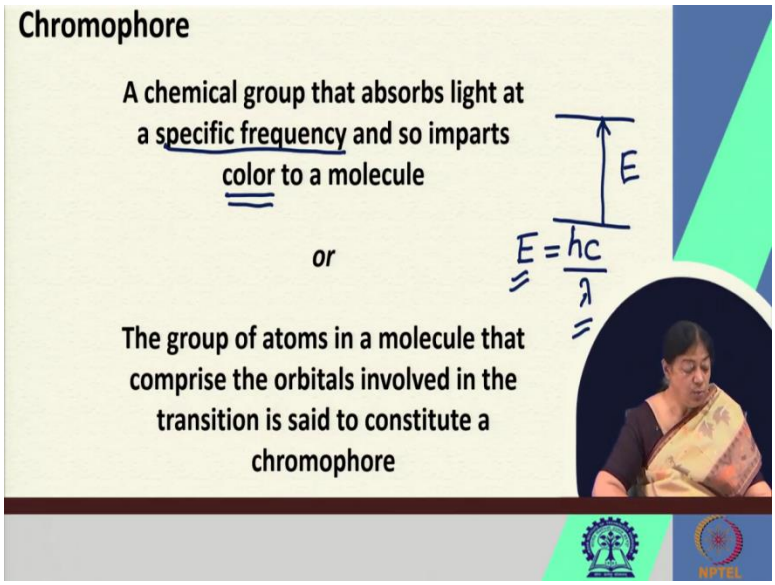
We also looked at the hydrophobicity in terms of a hydropathy index; we were looking at a specific way to demonstrate how we can find out a hydropathic or hydrophobic region of a protein. So, these are interesting properties of amino acids and they go in a very long way to understand how proteins behave, because these are the building blocks of proteins.

In this lecture, we are going to look at the spectroscopic properties of amino acids. Now, what do we mean by these spectroscopic properties of amino acids and what are the amino acids that are going to contribute to the spectroscopic characteristics?

(Refer Slide Time: 01:53)

So, when we look at the different keywords that we have looked at so far, we have not seen what chromophores are yet. So, what are these chromophores?

(Refer Slide Time: 02:02)



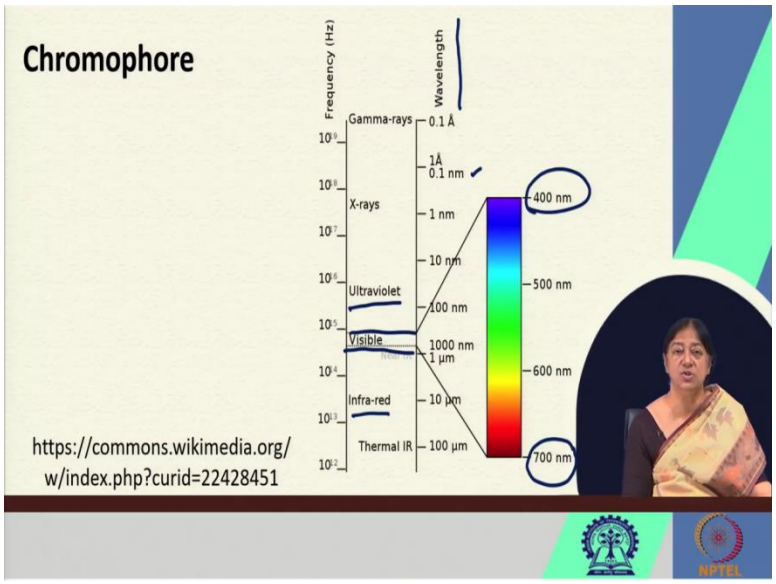
A chromophore is a chemical group that absorbs light at a specific frequency and it imparts color to the molecule. We can also define it as a specific group of atoms in a molecule that comprise the orbitals involved in the transition that constitute a chromophore.

So, what does this mean? In general terms, if we have two energy levels, we are going to have a transition and this transition is going to correspond to a specific energy. This energy, as we know, is related to the wavelength by the following expression:

$E=hc/\lambda$

So, when we look at this expression; we look at the wavelength and at the energy and we try to figure out, that if there were particular chromophores in proteins, what would the λ values be and where would they lie in the overall electromagnetic spectrum.

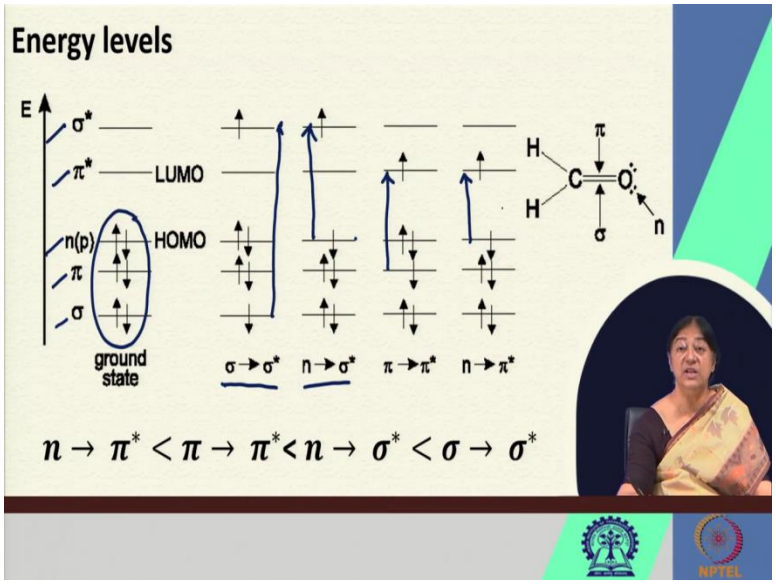
(Refer Slide Time: 03:19)



Now, what do we mean by this? When we look at the electromagnetic spectrum, we see that it has a range of wavelengths, which means it has a range of energy.

So, when we look at the specific wavelength values, we know that we have the visible region from 400 nm to 700 nm and we know that the colors we follow here are the VIBGYOR. Then we have the ultraviolet region (UV) that corresponds to a specific wavelength that is less than 400 nm (<400nm) and the infrared region (IR) is the region beyond 700 nm (>700nm). This means that when we look at the specific bonding or the specific electronic availability or the transition of these electrons, we have to figure out where these transitions are going to lie in terms of our electromagnetic spectrum.

(Refer Slide Time: 04:25)

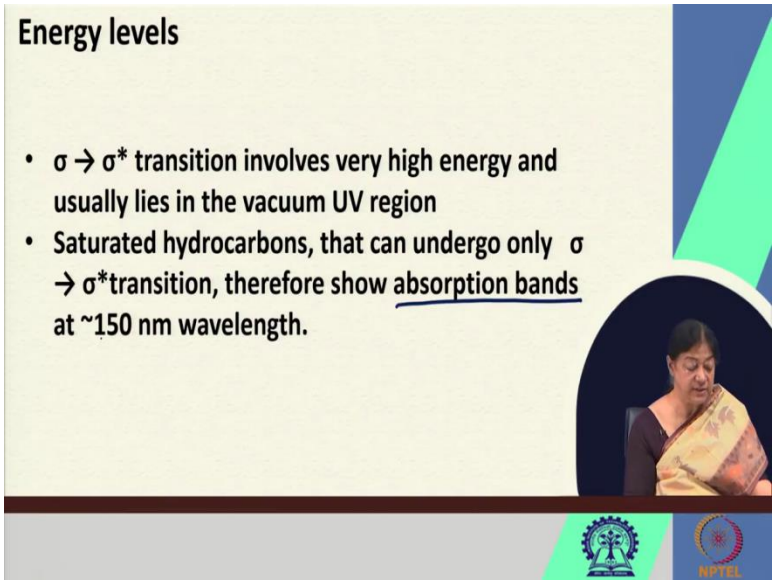


So, what are these energy levels? If we look at a generic form of these energy levels, we have the σ orbital, we have the π, we have the non-bonding, we have the π*and we have the σ*, in terms of increasing energy.

[Refer to slide] In the first transition here, from σ to σ*, we see that in the ground state we have all the electrons sitting in the σ, π and non-bonding. What is going to happen? A σ to σ* transition means that in the energy level, there will be an energy jump. Similarly, when we are going from n to σ*; we are losing one from n(p) and we are going up to σ*. In a π to π* transition, we are going from π up to π*; and similarly in a n to π* transition we will go up from n(p) to π*. Now, we realize that each of these transitions involve different energy values, which means that each of them belong to a different wavelength value.

This wavelength value is important into understanding where in the electromagnetic spectrum it falls; whether it's in the visible region, IR region, or the UV region and how this can be measured. So, if we look at the transitions, we see that the n → π* < π → π* < n → σ* < σ → σ* in terms of the value or in terms of the energy levels.

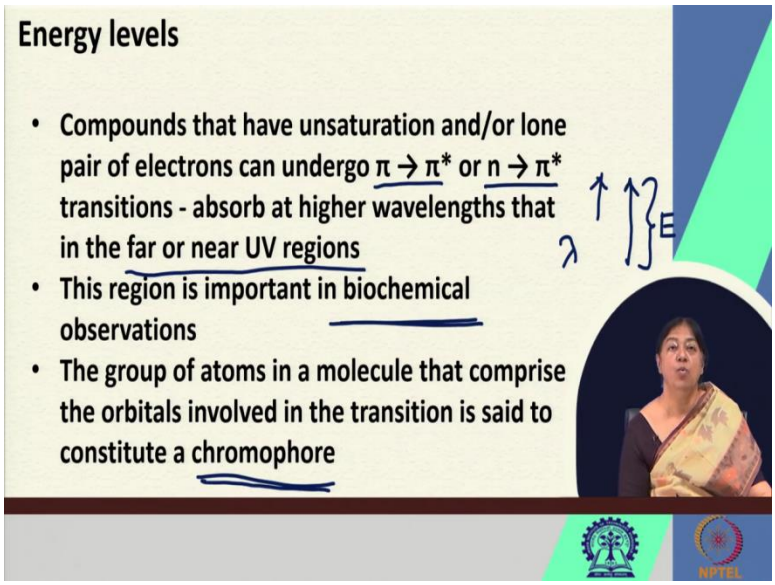
(Refer Slide Time: 06:29)



Now, if we look at the energy levels that we saw, we saw that the σ to σ*transition involves very high energy and usually lies in the vacuum UV region. Now, how was that known? Because from the energy value, we can calculate the λ and from the λ value, we can determine which region of the electromagnetic spectrum this falls into.

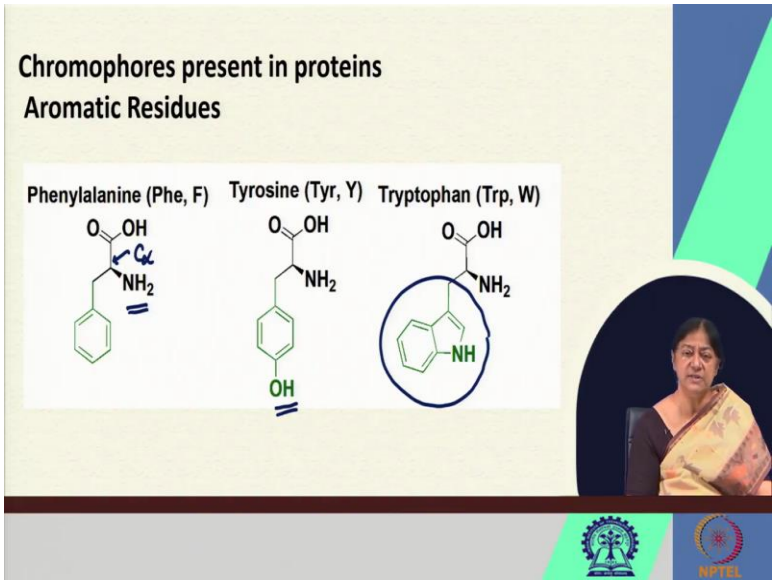
So, we have saturated hydrocarbons that undergo σ to σ*transitions only and show absorption bands in the 150 nm wavelength.

(Refer Slide Time: 07:03)



Similarly, when we look at compounds that have an unsaturation, i.e. they have double bonds in them or even a lone pair of electrons, they undergo π to π* or n to π* transitions. They absorb at higher wavelengths that are there in the far or near UV regions. So, what does this mean? This means that the value of the transition that we are looking at is from n to π* or from π to π*. The energy value that we are looking at here, has a particular λ value that is present in the far or near UV region. This region is important in all our biochemical observations. Why? Because, a group of atoms in a molecule consisting of orbitals undergoing transitions, would constitute of a chromophore and we know that studying this chromophore will help us understand what the chromophores are in proteins.

(Refer Slide Time: 08:18)

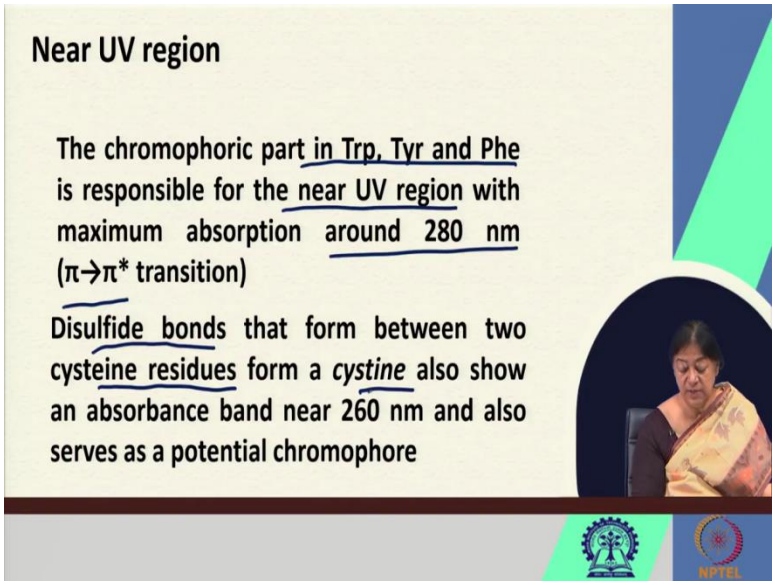


So, when we look at the chromophores in proteins, there are the aromatic residues(side chains) like - the phenylalanine, the tyrosine and the tryptophan, that are present in proteins. So, just like we had looked at all the representations of the amino acids, we know that there is a Cα carbon atom, which has the NH_2 group and the COOH group attached to it.

This is just a common representation. We know that when we look at the zwitterionic form, the Cα carbon atom is going to be attached to NH_3^+ and COO^- instead of the NH_2 group and the COOH group. [Refer to slide] Then we have tyrosine here, which has an OH group attached to it and we also have tryptophan, which contains an indole group.

These aromatic residues are the chromophoric moieties in proteins that bring about the spectroscopic characteristics in the proteins to a large extent, making them useful for an understanding of their concentration, their presence and so on.

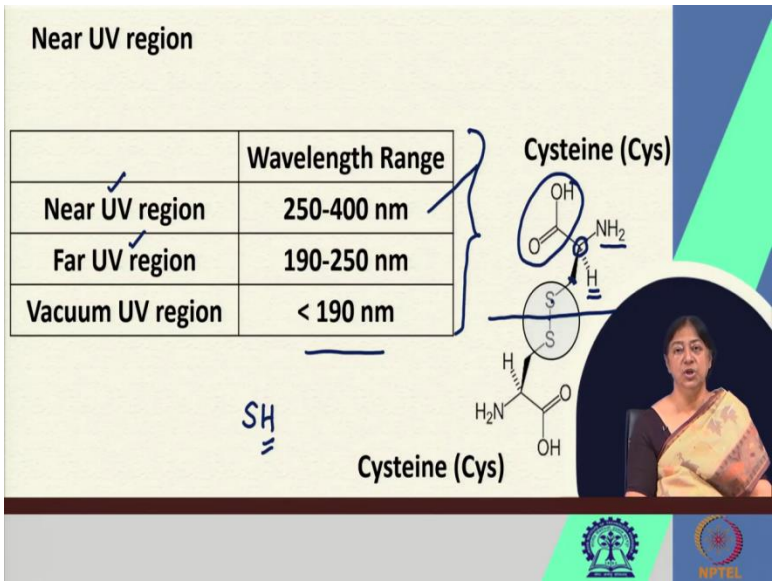
(Refer Slide Time: 09:39)



So, we look at the near UV region. What do we mean by the near UV region? We know that the visible region starts from 400 nm to 700 nm. So, the near UV region is close to 300-320 nm. So, the chromophoric part in tryptophan, tyrosine and phenylalanine is responsible for the near UV region with a maximum absorption around 280 nm which corresponds to the π to π* transition. So, it is important to understand that we have the three aromatic amino acid residues, that are going to contribute as chromophores in proteins and they are going to give a maximum absorption for the π to π* transition, that is going to fall in the region around 280 nm, which is the near UV region.

[Refer to next slide] The disulfide bonds that form between the two cysteine residues can form what is called cystine. They have an absorption around 260 nm and they can also serve as a potential chromophore; because 260 nm falls in the near UV region as well.

(Refer Slide Time: 11:11)

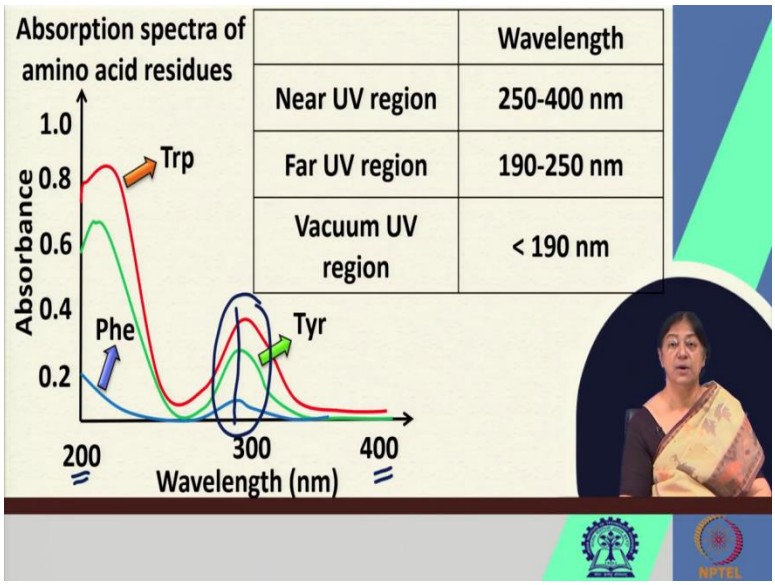


So, when we are looking at the near UV region, we have the wavelength range from 250 nm to 400 nm; we know that from 400nm, the visible region starts. So, we see that close to the visible region, we have the near UV region, followed by the far UV region from 190 nm to 250 nm. This is also important in biochemical techniques in terms of characterization of proteins and we have the vacuum UV region which is less than 190 nm.

The wavelength range that we see is going to be dependent on the energy levels. So, what are the energy levels dependent on? They are going to be dependent on the specific transition that we are talking about. So, whether we talk of an n to π* transition or a π to π* transition; this difference in the energy levels corresponds to a particular wavelength and if this wavelength falls within these mentioned regions, we will be able to observe these transitions in a UV spectrophotometer.

Now, this is where we take the example of cysteine. So, the amino acid cysteine consists of CH_2SH as the side chain. [Refer to slide] So, in the cysteine, we can easily identify the α carbon**,** the hydrogen, the amino part and the acid part associated with the α carbon. Apart from the Cα, there is also a Cβ atom, which is attached to a sulfur, which in turn is attached to a hydrogen, forming a CH_2SHside chain altogether. Now, what happens when the hydrogen is lost during the formation of the disulfide linkage? It leads to a particular moiety that is also going to absorb around 260 nanometers also contributing to [Refer Time: 13:28] a potential chromophoric moiety.

(Refer Slide Time: 13:34)

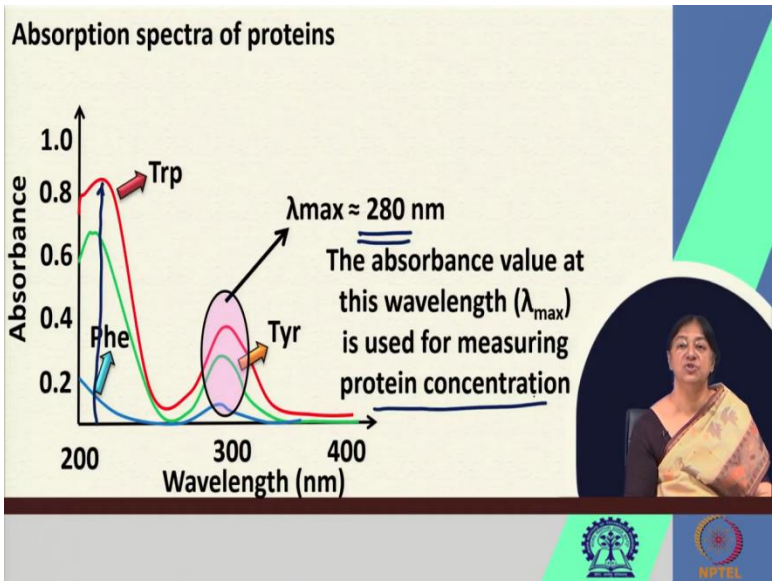


Now, coming to the topic of the absorption spectra of the amino acid residues. What do we mean by this absorption spectra? If we take the amino acid tryptophan or tyrosine or phenylalanine independently in what is called a cuvette for the machine known as the UV spectrophotometer and we scan a particular wavelength.

So, we are scanning from 200 nm to 400 nm. So, when we scan from 200 nm to 400 nm, we are scanning around the near UV region. Why do we want to do that? We want to see whether we get any specific absorption due to the scanning. What do we see? Yes, we do have absorption associated with the scanning from 200 nm to 400 nm and we see an absorption around this region [refer to slide].

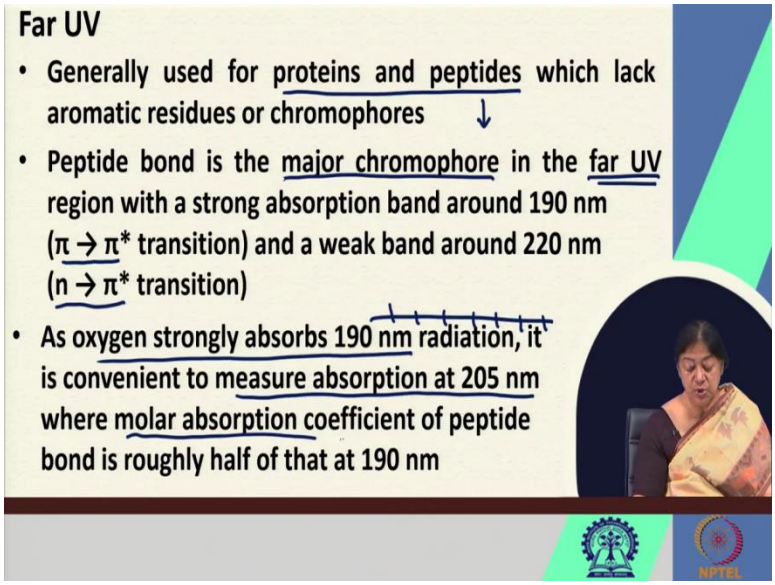
So, what do we mean by the absorption? We mean that we have peaks and these peaks confirm that tryptophan, tyrosine or/and phenylalanine is present. And this gives us an idea of where to look at to see whether or not we have a protein present. If the protein solution or the protein of interests has any of these aromatic amino acid residues present in them, which is usually the case; then we will see a presence of a peak around the region of 280 nm indicating that this is a protein having this amino acid.

(Refer Slide Time: 15:15)



So, here the λ_{max} is around 280 nm indicating that this solution does have a protein or does have an amino acid present in it, that is going to correspond to 280 nm. Here we also see some transitions following at a lower wavelength, meaning a higher energy, meaning a different transition. However, we are interested in this particular transition of 280 nm, because the absorbance value of this particular wavelength is used to measure the protein concentration, which we will see in a moment.

(Refer Slide Time: 16:03)



So, the far UV region is usually used for proteins and peptides that do not have aromatic residues or chromophores. This is because when we are talking about a chromophoric unit, the presence of the aromatic amino acid residues - the tryptophan, tyrosine, phenylalanine - are going to give

specific transitions at 280 nm, that will make it very useful to identify whether it is a protein solution or whether a particular peptide is present or not.

But what happens if a particular peptide does not have any of these aromatic amino acid residues? Then we have to resort to looking at the far UV region to identify whether we have a particular peptide bond. Now, the peptide bond is again a major chromophore in the far UV region.

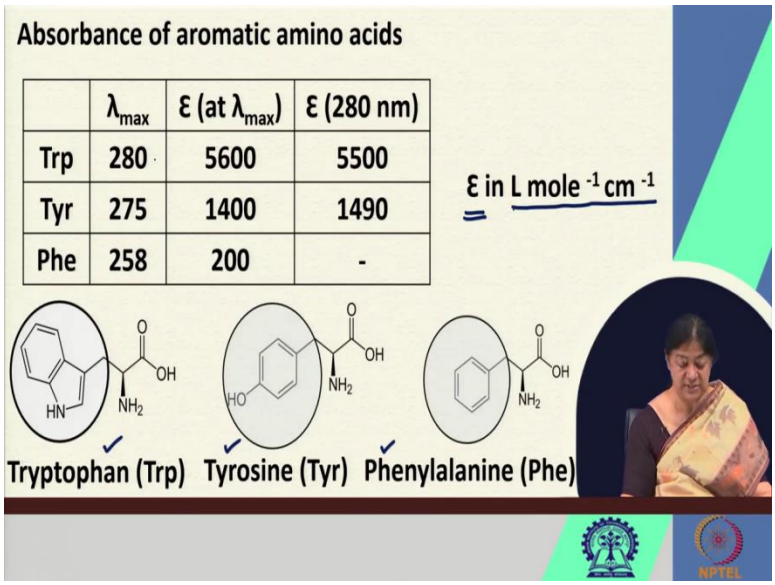
Why? Because in this particular case, we are also going to get what is called a π to π^* transition and an n to π^* transition. For the π to π^* transition, there is a strong absorption band around 190 nm whereas there is a weak band around 220 nm for the n to π^* transition.

Now, these transitions occur for the peptide bonds; all of the proteins. In the next lecture, when we will study the interactions and the formation of the peptide bond, we will see that this peptide bond is present throughout. So, the identification of the peptide bond is going to mean that, with a lot of these peptide bonds that are present; we are going to have a large absorption band at 190 nm and at 220 nm.

So, for the identification of proteins we usually resort to the aromatic amino acid residues - tryptophan, tyrosine or phenylalanine; because these are relatively less abundant in proteins than the peptide linkage, which is present definitely in proteins, because they link the amino acids together.

Now, a problem here is that oxygen also strongly absorbs at 190 nm. So, we usually measure the absorption at 205 nm, where the molar absorption coefficient of the peptide is actually half of that observed at 190 nm, which is the strong absorption band. So, ideally when we are looking at a peptide that does not have an aromatic amino acid present; then it is preferable to measure the absorbents at 205 nm rather than 190 nm, despite the fact that the strong absorption band occurs at 190 nm. Why? This is because oxygen also strongly absorbs at this point and it will interfere with our measurement.

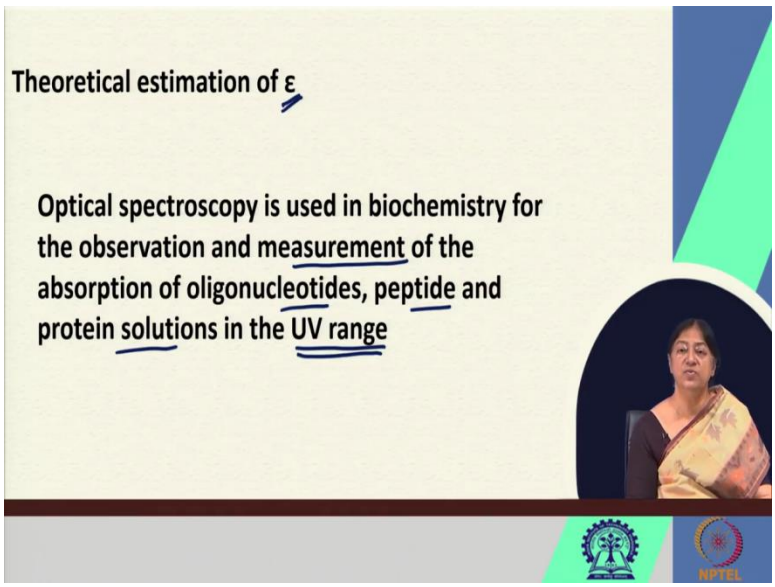
(Refer Slide Time: 19:27)



So, if we look at the absorbance of the aromatic amino acids, we look at what is called an extinction coefficient that gives us a measure of the capability of the particular moiety or the particular chromophore in this case, to absorb light. Now, when we look at the extinction values here; we know that these are the side chains of tryptophan, tyrosine and phenylalanine.

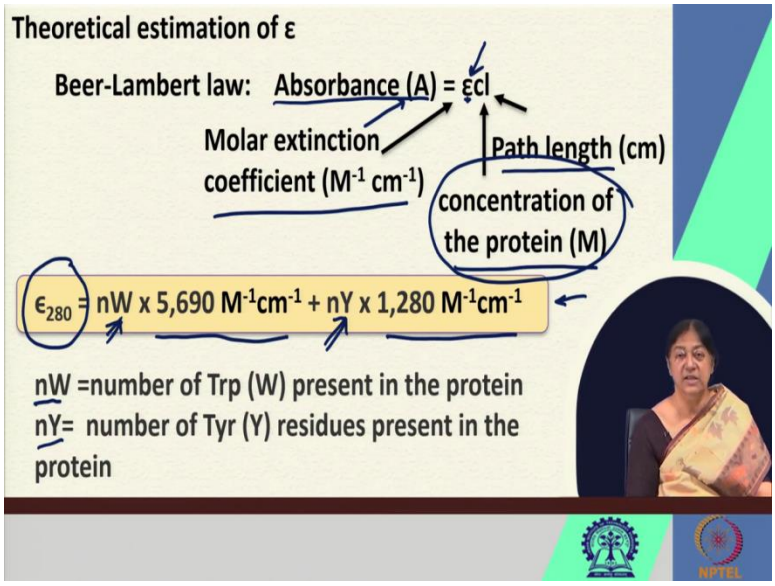
The extinction value is given in the units of L mole^{-1} cm^{-1}. Now, what is this or what are the λ_{max} values? The λ_{max} values are the particular spectrum of the amino acid listed here and we look at this value corresponding to where the peak is observed in the particular spectra that is taken for the amino acids.

(Refer Slide Time: 20:26)



Now, when we look at the theoretical estimation of this ϵ value; then we are looking at a very important aspect in trying to understand how this spectroscopic characteristic is used in biochemistry for the observation and the measurement of the absorption of oligonucleotides, peptides and protein solutions all in the UV range. Why do they fall in the UV range? This is because all the transitions that we observe from n to π^* and π to π^*, the energy transitions or the difference in the energy levels falls in the UV range for the particular wavelength.

(Refer Slide Time: 21:17)



The Beer-Lambert's law, as you know, gives us the absorbance value. From this absorbance value, we can, from a measure, determine the molar extinction coefficient of a specific solution; given that we know the concentration of the protein that has been taken as well as the path length, which is usually 1cm for the cuvette that is used in the machine.

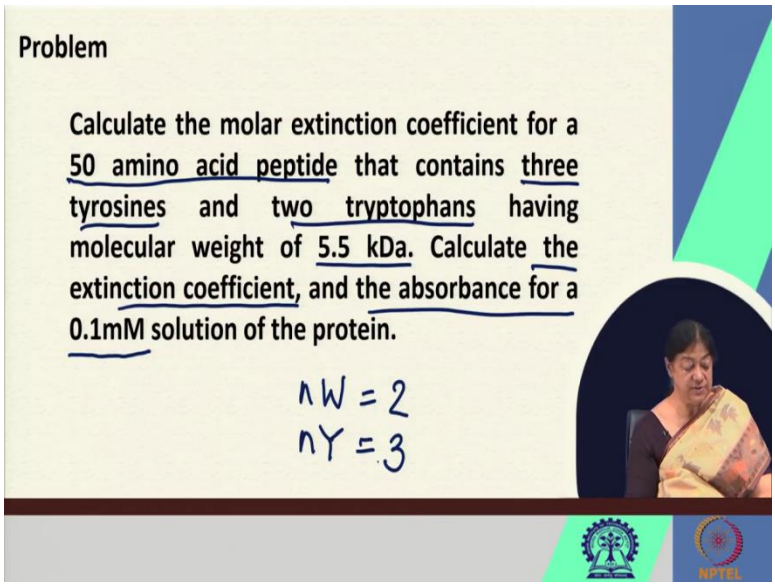
So, this is important when we are trying to understand a theoretical estimation to figure out what the extinction coefficient of a particular protein is going to be at 280 nm, given that we know the number of tryptophan is nW and that of tyrosine is nY. W is the one letter code for tryptophan and Y is the one letter code for tyrosine.

Now, given the value that we had for the extinction coefficient for the specific amino acids, these are additive values. So, if we have an nW, the number of tryptophan residues that are present in the protein and nY, the number of tyrosine residues in the protein, we can get an estimate of what the value of this extinction coefficient is going to be, given that we know the number of tryptophan and the number of tyrosine.

Now, given that we know or we can determine, what the absorbance value is from the machine when we find out the OD or the optical density value from the machine (Refer Time: 23:20). And then given the extinction coefficient that we are calculating from a knowledge of the number of tyrosine and tryptophan, what we can do you realize is, we can find out the concentration of the protein; a very very important factor, a very very important calculation, a

very common calculation that is done in any protein chemistry laboratory, because you have to determine the protein concentration in any experimental observation that you are going to make.

(Refer Slide Time: 23:55)

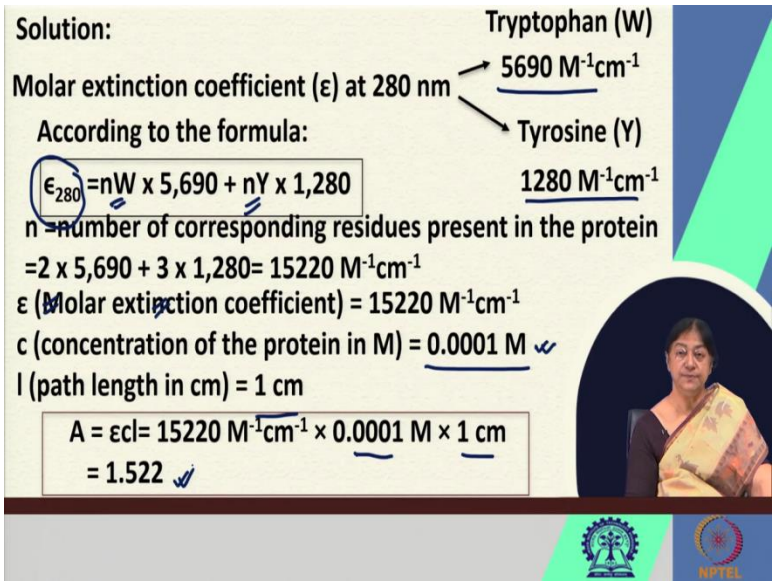


So, let us look at a simple problem based on this. So, the problem states that, we have to calculate the molar extinction coefficient for a 50 amino acid peptide. Now, what is this amino acid peptide? Now, we have to look into 50 amino acid peptide that contains three tyrosines and two tryptophans and it has a molecular weight of 5.5 kDa. So, there are certain new terms that we have we are learning here; one is the kDa. Now, what is the kDa?

The kDa tells you, a Da is 1 unit of molecular weight; 5.5 means 5500 kDa. So, when the question is that, we have to calculate the extinction coefficient and the absorbance for a 0.1 mM solution of the protein. So, the first thing to try and find out what the extinction coefficient is means, we need to know n; we need to know what is nW and we need to know what is nY.

So, nW we know is equal to 2, because there are 2 tryptophans and nY is equal to 3, because there are 3 tyrosines.

(Refer Slide Time: 25:27)

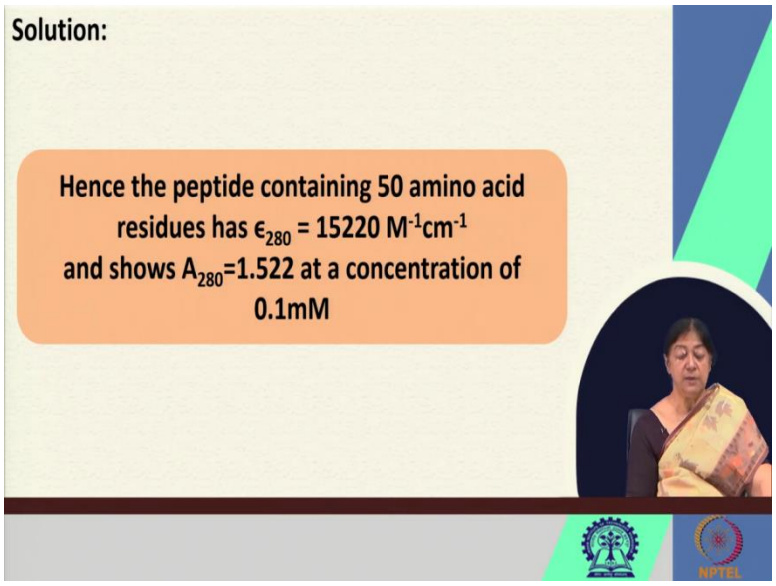


So, the molar extinction coefficient for tryptophan is this, for tyrosine is this value. Then according to the formula that we looked at previously; we have the value of nW here we have the value of nY here, so we can find out the extinction coefficient for this particular set or this particular peptide. So, the number of corresponding residues is 2 as we learn 2 tryptophan 3 tyrosine and there is our value.

So, once we have the molar extinction coefficient, we know the concentration, we know the path length; we have to take the concentration, it was given as 0.1 mM. So, we have to be very careful that we, because our units have to cancel out absorbance values are unit less.

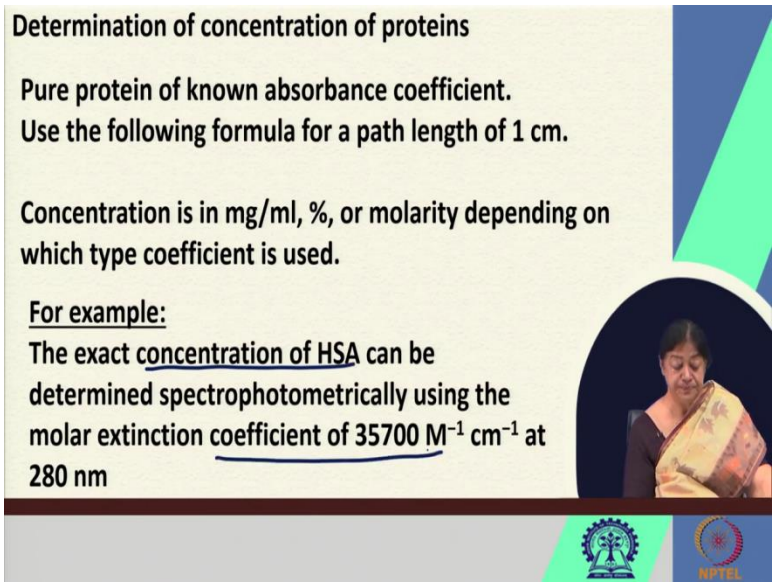
So, we have an absorbance value 15220; where did we get this from? We got this from an understanding of the tryptophan and the tyrosine and the concentration, the path length giving us an absorbance value. So, that is the absorbance value.

(Refer Slide Time: 26:39)



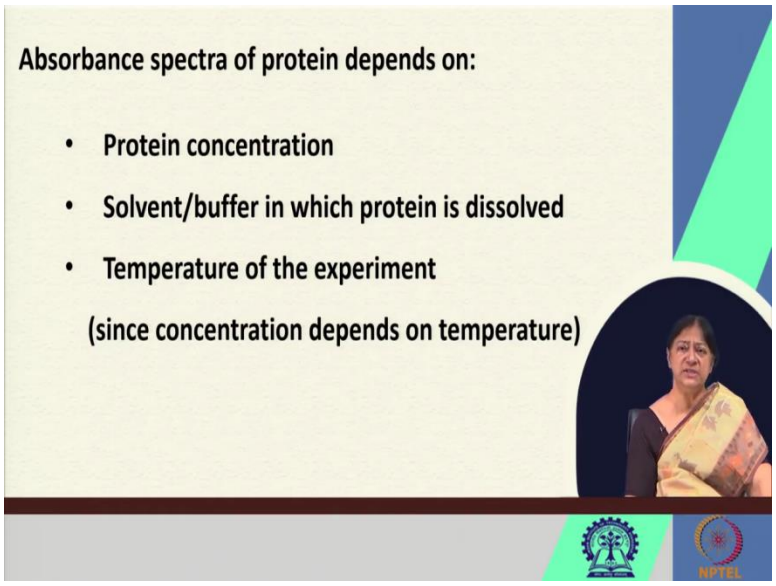
So, the peptide containing 50 amino acid residues has a specific extinction coefficient corresponding to 15220 $M^{-1}cm^{-1}$ and an absorbance value of 1.522.

(Refer Slide Time: 26:59)

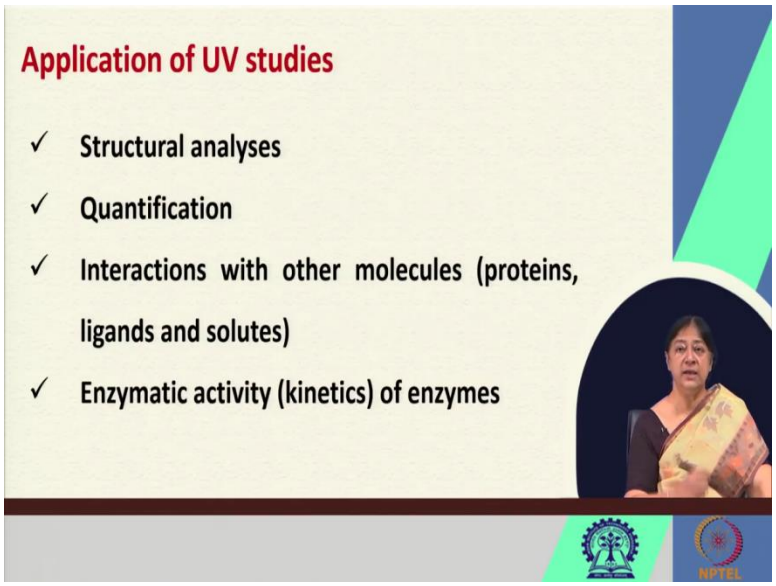


This is used for the determination of the concentration of proteins as we mentioned and we can calculate the exact concentration of many many proteins, given we know the extinction coefficient.

(Refer Slide Time: 27:15)

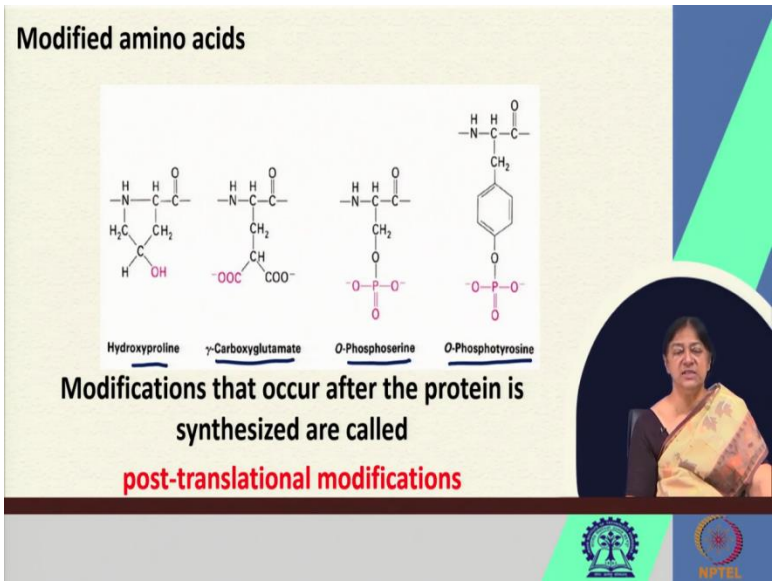


(Refer Slide Time: 27:22)



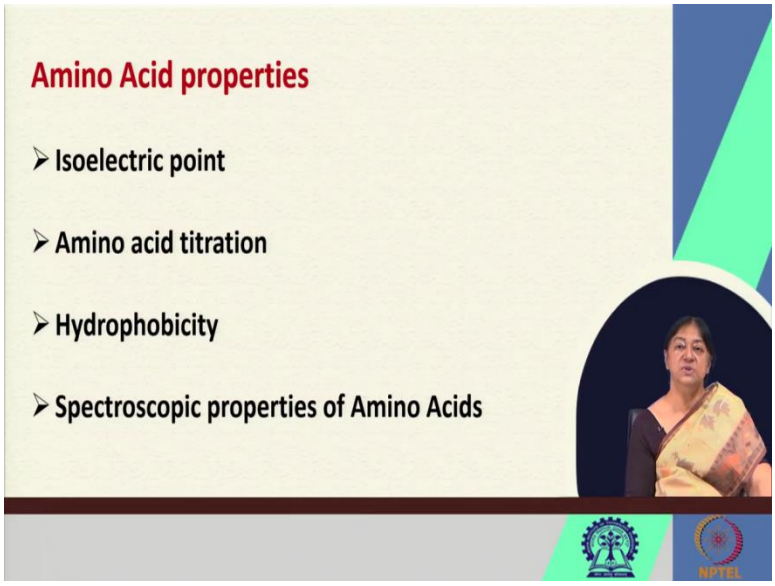
So, the absorption spectra depends upon the protein concentration, solvent buffer, the temperature, which we will see as we go along in an understanding about how we can apply the UV spectroscopic studies to proteins in general.

(Refer Slide Time: 27:28)

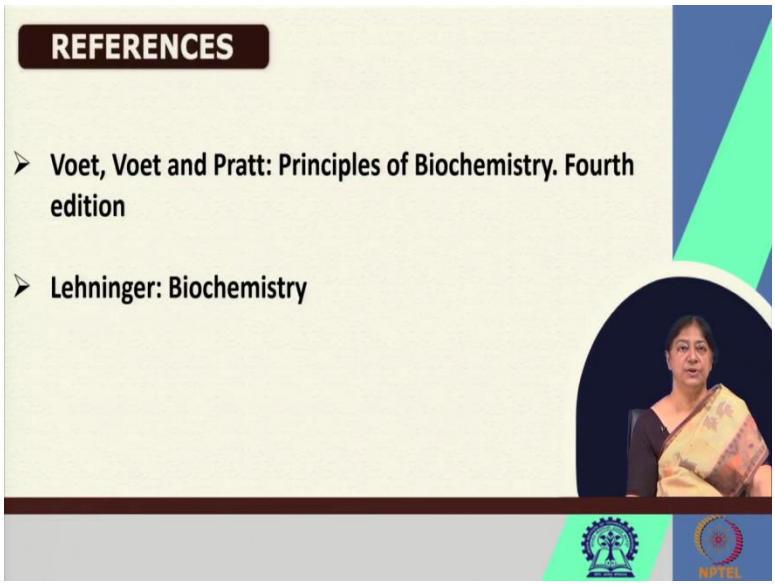


Another small thing that needs to be covered or just an understanding, that there are some modified amino acids; that are just as you can see the names hydroxyproline, carboxyglutamate, phosphoserine, and phosphotyrosine. And there are other combinations that are called post translational modifications that occurred after a protein is synthesized, where we have a combination of amino acids that sometimes give rise to a different moiety all together.

(Refer Slide Time: 28:03)



(Refer Slide Time: 28:26)



So, the amino acid properties that we studied were the isoelectric point, the amino acid titration, the hydrophobicity, and spectroscopic properties of the amino acids that give us an understanding of amino acids in general. Our next lecture will be on the combination of amino acids now in the formation of a peptide bond, before we embark on an understanding of protein structure.

Thank you.