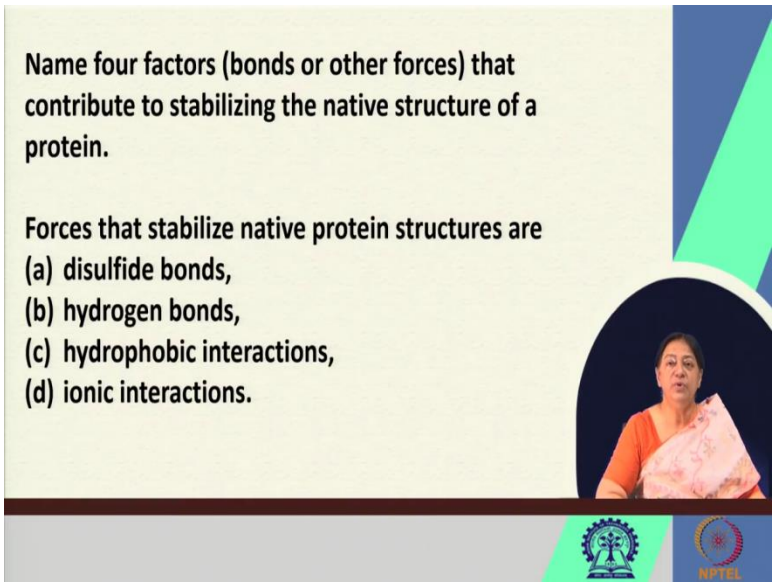
Fundamentals of Protein Chemistry
Prof. Swagata Dasgupta
Department of Chemistry
Indian Institute of Technology, Kharagpur

Module - 03
Protein Structure and Stability
Lecture - 15
Discussion Class

In our discussion class for module 3 based on protein structure and stability, we will revisit some of the topics we studied and a few of the problems associated with these topics.

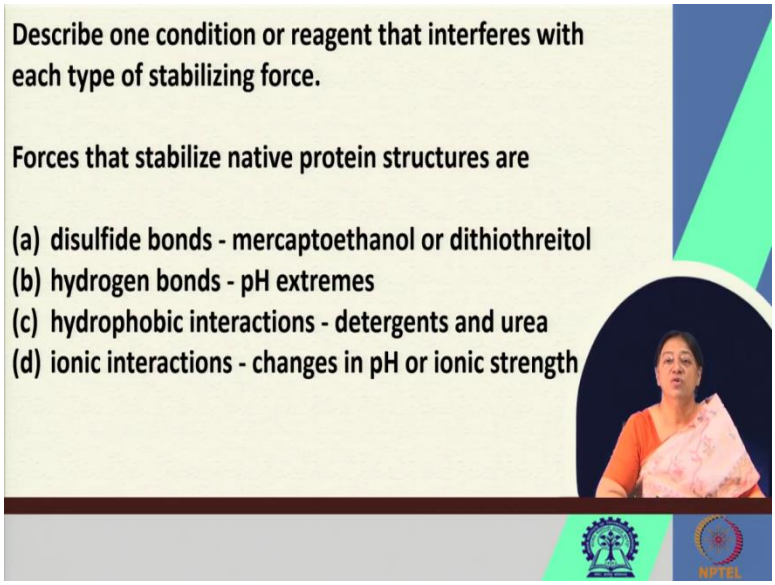
(Refer Slide Time: 00:29)



So the first thing that we need to look at, are the factors or the bonds or forces that contribute to stabilizing the native structure of the protein. These forces or these factors will come again when we consider protein-ligand interactions protein-protein interactions. And the forces that stabilize the native protein structures; the covalent ones, the non covalent ones are the ones that we will be seeing in other types of interactions as well.

For the basic forces that we know that stabilize these native protein structures; we have the disulfide bonds, the other covalent linkage apart from the peptide bond, we have the hydrogen bonds, we have the hydrophobic interactions and we have the ionic interactions.

(Refer Slide Time: 01:15)

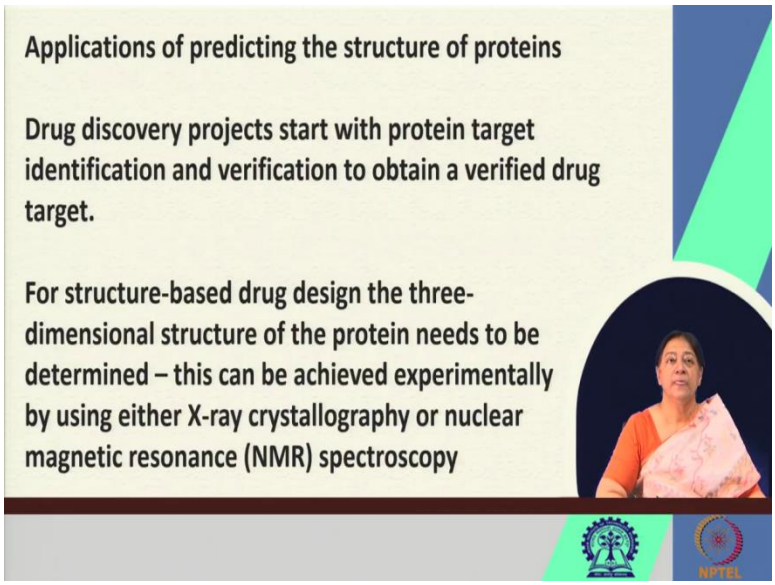


If we were to look at forces or a condition or a reagent that would interfere with the particular type of stabilizing force that has been mentioned, we will have to go back to the lecture on denaturation of proteins, to understand which component or which particular reagent could break the specific type of interaction or rather interfere with that specific type of interaction.

We look at the disulfide bonds and we know it is β-mercaptoethanol or dithiothreitol. Hydrogen bonds would be pH extremes that would affect the protonation, deprotonation of the ionic side chains. We have the hydrophobic interactions, where we know the interference can be with detergents and urea. We will have a look at mercaptoethanol and a specific detergent later on in this lecture.

And for the ionic interactions again, it would be the changes in the pH or the ionic strength that are going to affect the overall electrostatic interactions.

(Refer Slide Time: 02:33)

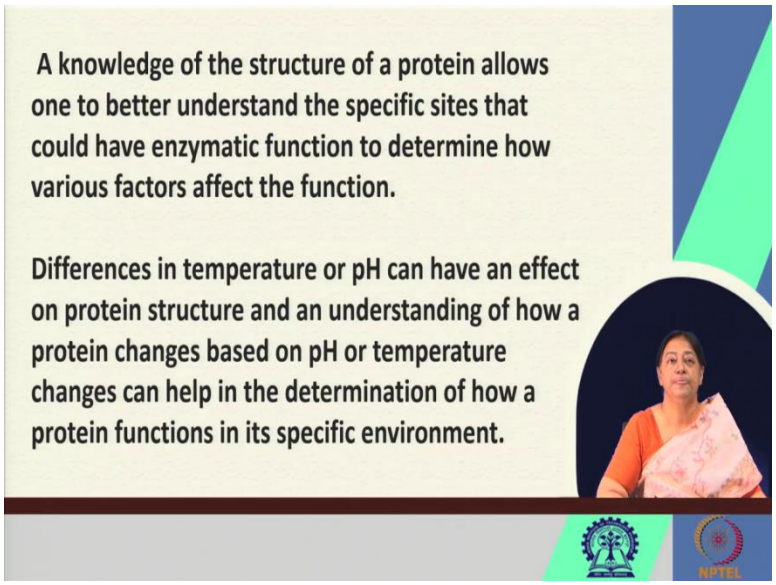


The understanding of the structure of proteins or a knowledge of the structure is increasingly important. Why, because these drug discovery projects have a protein as a target. And it is required to have a verified drug target, so that we know a receptor molecule, that the drug can be designed against; in what is called rational drug design.

For structure based drug design, the three dimensional structure of the protein, is extremely important. We understood the importance of structure and function and essentially this can be achieved experimentally by using either X-ray crystallography or nuclear magnetic resonance NMR spectrometry.

The reason that we consider these specific structure function relationships, is because we want to develop a specific drug against a specific protein, to inhibit its action or to look at changes in protein-protein or protein-ligand or protein-drug interactions and how they can be brought about by a knowledge of the structure. We looked a bit at protein structure prediction methods, but in a very small scale.

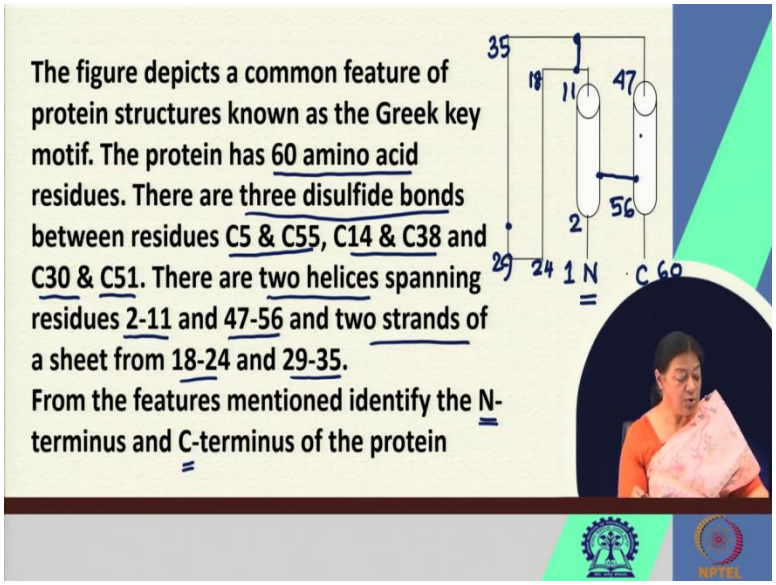
(Refer Slide Time: 04:10)



A knowledge of this structure of a protein, allows one to understand the specific sites that could have say enzymatic function and to determine how these various factors would affect the function.

Differences in temperature pH can also have an effect on protein structure as we saw. And an understanding of how the protein structure changes can occur, based on these pH or temperature changes, can help in the determination of how this protein function is going to be affected in this specific environment.

(Refer Slide Time: 04:52)



Now we have a problem associated with disulfide linkages and a protein structural motif. Here [refer to slide] there is the protein structure motive known as the Greek key motif. There are certain information given about this protein. This protein has 60 amino acid residues, there are 3

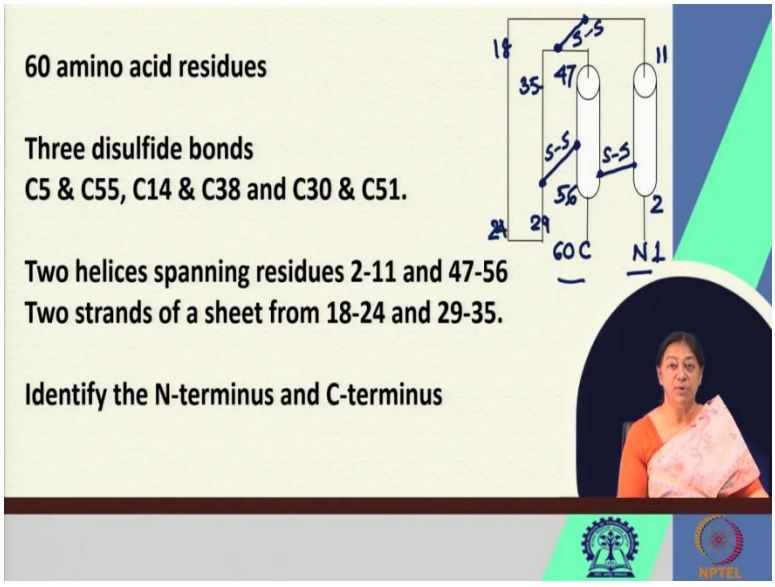
disulfide bonds between C5 & C55, C14 & C38, C30 & C51. As we see the diagram given on the right, there are two helices that span residues 2 to 11, 47 to 56 and two strands of the former β-sheet that are from 18 to 24 and 29 to 35.

Given this information what we have to do is, we have to identify the N terminus and the C terminus of the protein. If we consider this [refer to slide] to be the N terminus and this to be the C terminus, and given the information we have to see that we have the first disulfide linkage between C5 and C55, cysteine 5 and cysteine 55. So this has to be residue number one and this being C, this has to be residue number 60 because it is a 60 amino acid protein. We can consider that we have the helices spanning from 2 to 11. We have 2 to 11 that is one helix and the other is from 47 to say around 56. And what do we have? We have a disulfide linkage from say 5 to 55. So this could be the structure.

Then, we have our two strands from 18 to 24 and again from 29 to around 35, this is just a rough estimation. The next disulfide linkage we have, is between 14 and 38. If this is 11 [refer to slide], so 14 is somewhere here and 38 can be somewhere here. This would be another disulfide linkage.

Then our third disulfide linkage is between 30 and 51. 30 is here and 51 is somewhere here. Which means that the assumption that we took initially, that this is the N and this is the C terminus is not correct.

(Refer Slide Time: 08:00)



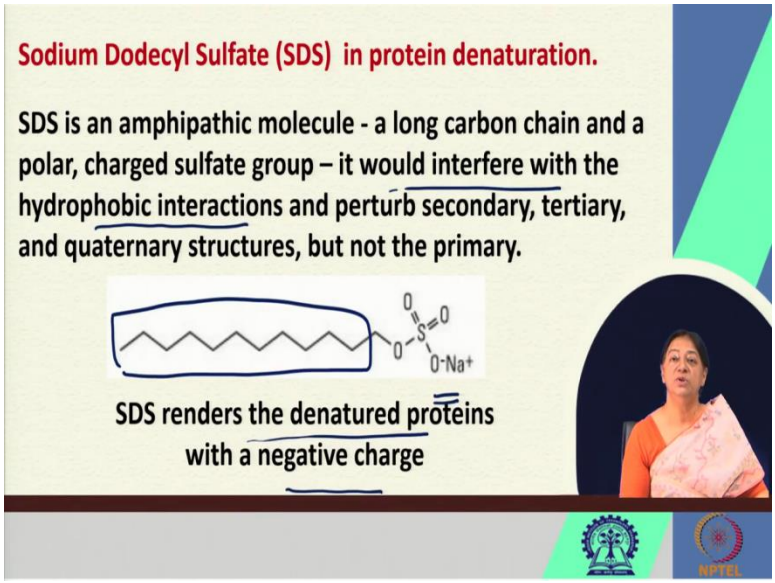
Let us see how we can work this out. Now consider this as the N terminus residue 1, this as the C terminus residue 60, these are the only two possibilities that we have. Again we have our helix spanning from 2 to 11, the other one is from 47 to 56. And we have our two strands going from 18 to around 24 and then 29 to somewhere around 35.

If we look at our information again regarding the three disulfide bonds, we have from 5 to 55. So, 5 is somewhere here and 55, so here is one of the disulfide linkages. Our next would be 14 to 38. 14 would be somewhere here and 38 would be somewhere here.

We would have one this way and we have another one from 30 to 51. So 30 would be somewhere here and 51 somewhere here. These are the three disulfide linkages the S – S linkages. The two options that we looked at the previous option where we considered this one to be the N terminus and that to be the C terminus did not work out, because we could not account for one of the disulfide linkages.

However, when we consider this to be the N terminus and this to be the C terminus we were allowed or permitted to explain all the properties that we had.

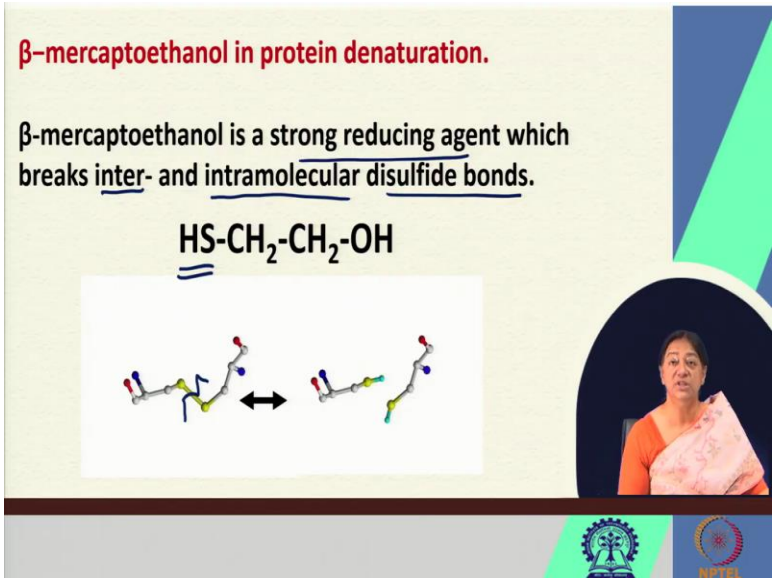
(Refer Slide Time: 09:56)



A detergent called sodium dodecyl sulfate, SDS is commonly used in protein denaturation. It is an amphipathic molecule, a long carbon chain and has a polar charged sulfate group. It would in its formation interfere with the hydrophobic interactions because of its long hydrophobic tail and it can perturb the secondary tertiary and quaternary structures, but not the primary.

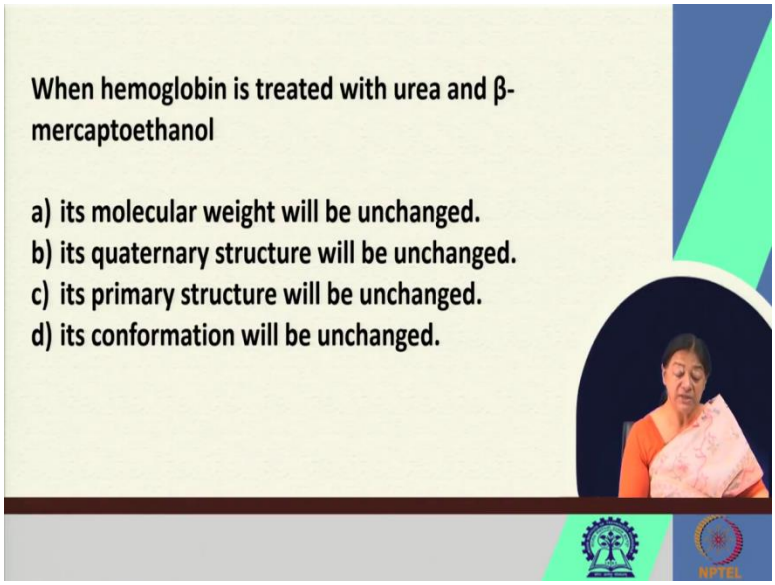
Interestingly, this renders the denatured proteins with the negative charge, something that is extremely important in the specific biophysical methodology that we used to determine the molecular weight of proteins called SDS-PAGE. Sodium dodecyl sulfate polyacrylamide gel electrophoresis which we will study when we do biophysical methods.

(Refer Slide Time: 10:55)



The β-mercaptoethanol, is a strong reducing agent, which breaks the inter and intramolecular disulfide bonds. So here [refer to slide] is a disulphide bond. Because of the thiol group here, this can break the disulphide bond, which is extremely important in considering a protein denaturation methodology.

(Refer Slide Time: 11:22)

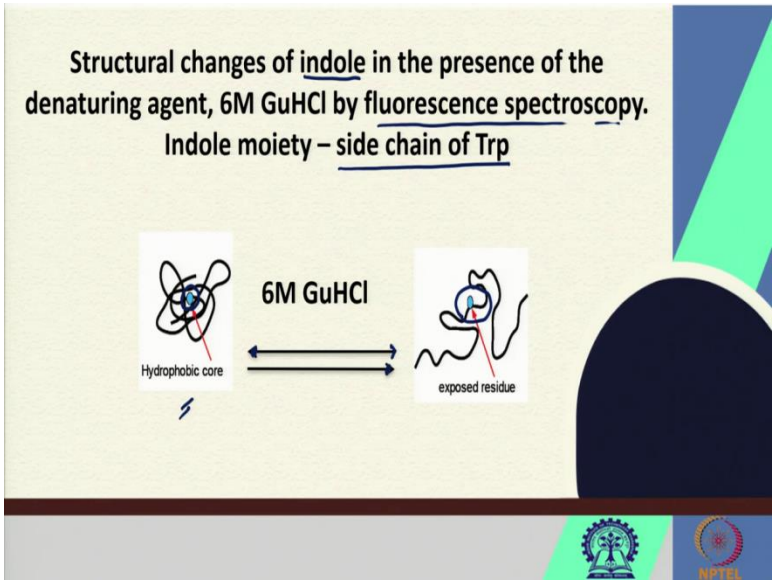


If we have a protein hemoglobin and we treat it with urea and β-mercaptoethanol, which of these is correct? Will the molecular weight, its quaternary structure, primary structure and conformation remain unchanged?

We know that when we are treating it with urea and β-mercaptoethanol, this is going to denature the protein. Hemoglobin being a tetrameric protein, is going to break up into its individual

monomeric units, the quaternary structure will change, the conformation will change. The molecular weight will be unchanged in totality and primary structure will also be unchanged.

(Refer Slide Time: 12:12)



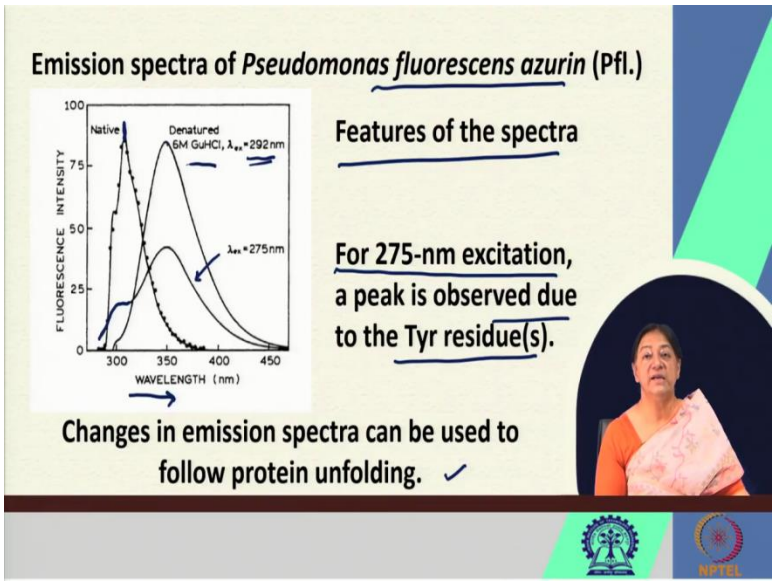
Now, when we look at structural changes of a specific indole moiety in the presence of a denaturating agent, such as 6M GuHCl, we mentioned in the lecture related to protein denaturation, that there were certain methods by which we could actually monitor the change. And in this case, we have a problem where the monitoring has been done by fluorescence spectroscopy.

What do we know about the indole moiety? The indole moiety is the side chain of tryptophan which shows fluorescence emission. Which are the three amino acids that are involved in UV absorbents mainly? The phenylalanine, the tyrosine and the tryptophan of which tryptophan has the highest extinction and it gives a high fluorescence emission as well.

So what happens is, if this particular tryptophan that is shown in a blue sphere here [refer to slide], is present in the hydrophobic core of the protein as we add the GuHCl, the arrow should be in this direction. What happens is, we have an unfolding and, due to that unfolding what happens, this will get exposed.

Now as it gets exposed what is going to happen? The signal is going to change or there is going to be some emission of the specific indole moiety.

(Refer Slide Time: 13:39)

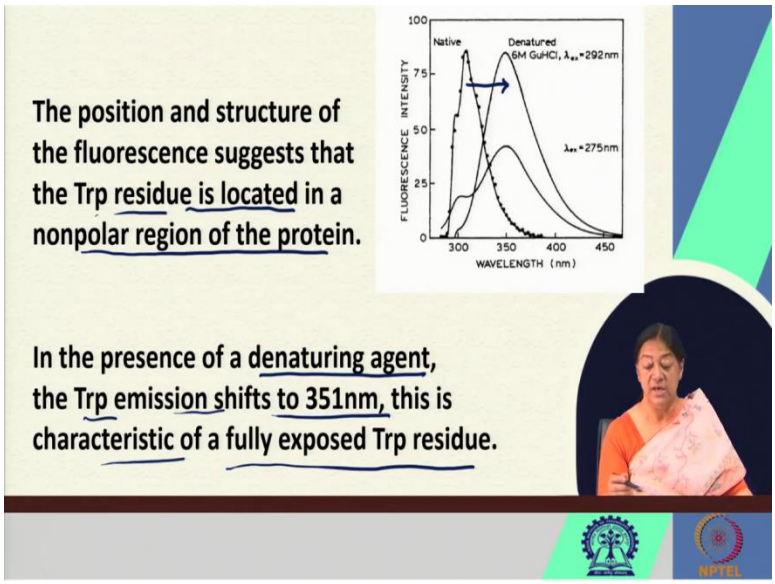


This [refer to slide] is the emission spectra of this particular type of protein azurin. What do we observe here? Here we have the fluorescence intensity versus the wavelength in the graph here. Here is our peak for the native and here we have what is called a red shift towards a longer wavelength for the denatured 6M GuHCl protein, indicating that these are the specific features of the spectra that we have to observe.

We see that for the native one we have an emission at a particular wavelength, for the denatured one we had a red shifted emission spectra. The changes in the emissions spectra now can be used to follow the protein unfolding. So what do we see? We see a red shift indicating that it has the tryptophan, in this case it is exposed. If we look at the excitation for 275 nm, this is the emission spectra that we see and this is for 292 nm excitation.

What we observe in this case is, we have another peak showing emission due to the presence of the tyrosine residues as well. So for tryptophan fluorescence emission, it is preferable that the excitation is done at around 295 nm.

(Refer Slide Time: 15:18)

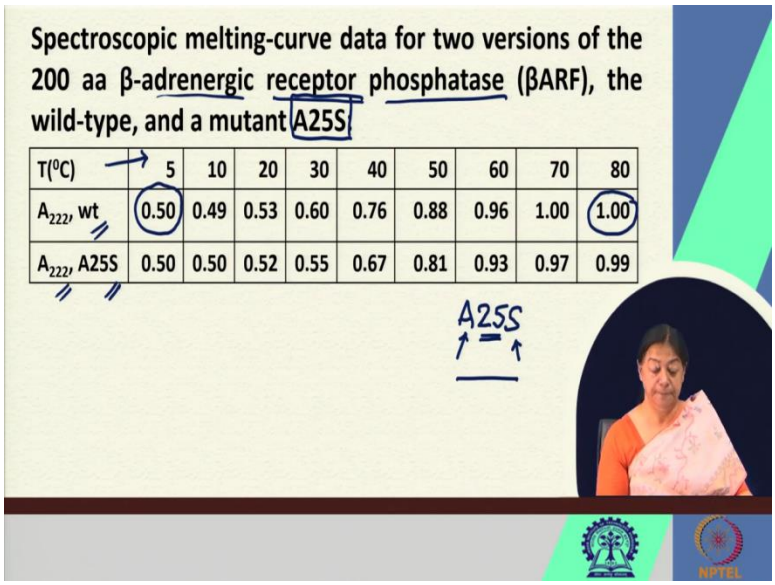


The position and the structure of the fluorescence spectra indicates that the tryptophan residue is located in a nonpolar region of the protein. Why - because from the native to the denature we have a red shift.

In the presence of the denaturating agent, in this case GuHCl, the tryptophan emission shifts to 351, which is characteristic of a fully exposed tryptophan residue. So just by looking at the fluorescence emission spectra, we can understand the positional preferences of the tryptophan and see where it had been located.

This is easier if you have say one tryptophan in your molecule, where you have the tryptophan either buried or exposed. If it is buried, as it is in this case and we denature the protein then what happens is, the tryptophan amino acid is exposed, giving us a fluorescence emission spectra such as seen here.

(Refer Slide Time: 16:23)

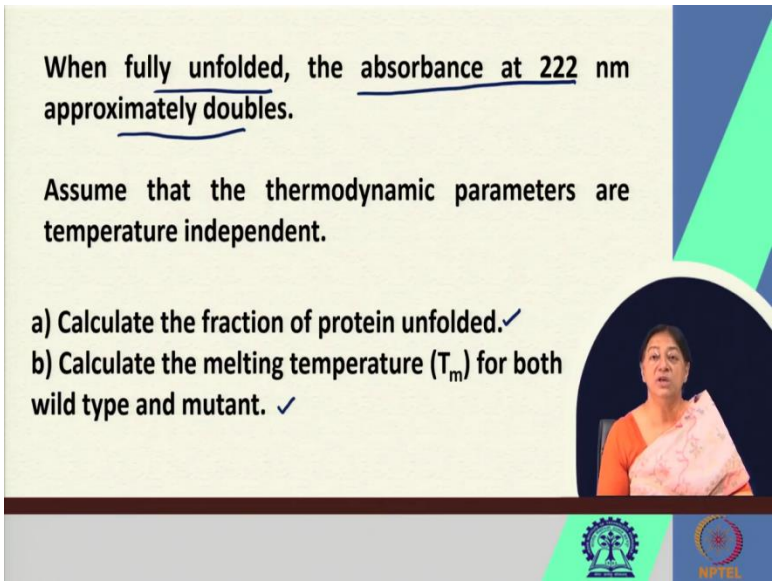


Another problem relating to the melting curve of two versions of an adrenogenic receptor phosphatase there is a wild type and a mutant. Now if we look at the way a mutant is expressed, we have it written as A25S what does this mean? This means A is alanine. The alanine at the 25th position has been mutated to serine. This is what this nomenclature means.

Let us now observe [refer to slide] the table. We have the temperature given to us at 5, 10, 20, 30, 40 up to 80 degree centigrade. We have the absorbance value at 222 nm, that is what the subscript means. For the wt meaning the wild type and we also have the same absorbance value for the mutant protein.

This is a specific experiment that has been conducted for the wild type protein and for the mutant protein, at different temperatures. What has been monitored? The absorbance at 222 has been monitored.

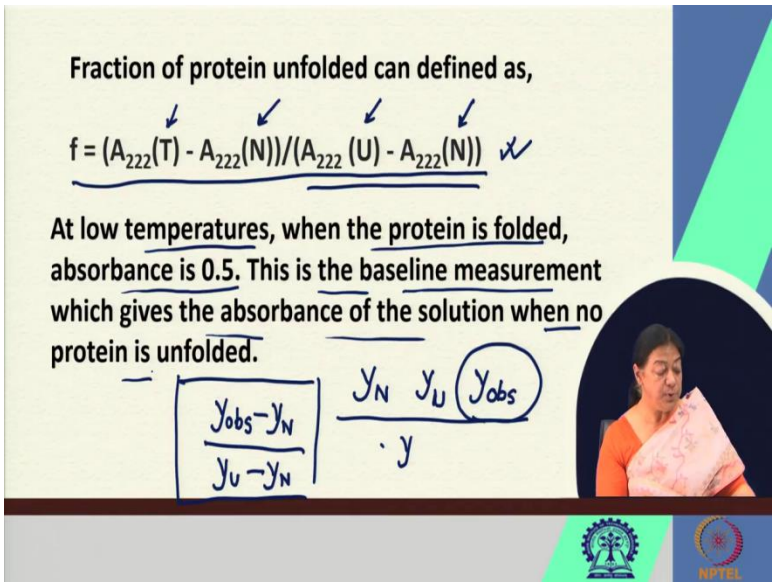
(Refer Slide Time: 17:50)



What is the information that we can get? When fully unfolded, the absorbance at 222 nm approximately doubles. Let us see the table. What we see at 5°C and what we see at 80°C; we see it almost doubling. So if we assume that the thermodynamic parameters are temperature independent, we can calculate the fraction of protein unfolded and we can also calculate the melting temperature for the wild type and the mutant.

Now how do we do that? What do we have? We have the information about the temperature and we have the information about the absorbance at 222 nm - both for the wild type and the mutant.

(Refer Slide Time: 18:44)



In this case our y is absorbance at 222 nm and we want to find the fraction of the unfolded part We have a measure for the native form, we have a measure for the unfolded form and we have other observations based on the intermediate values that we have.

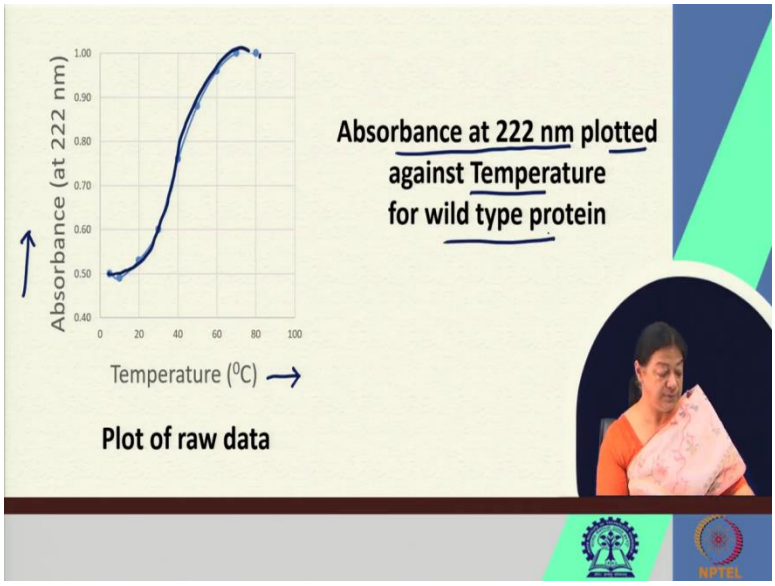
We have these values and from this we can calculate the fraction unfolded. This means that this [refer to slide] is the fraction. So we have: $y_{obs} - y_n / y_u - y_n$. This gives us a measure of the fraction of protein unfolded.

$f = (A_{222}(T) - A_{222}(N)) / (A_{222}(U) - A_{222}(N))$

So this denominator is actually a constant, because it is just the range of the values that we are looking at.

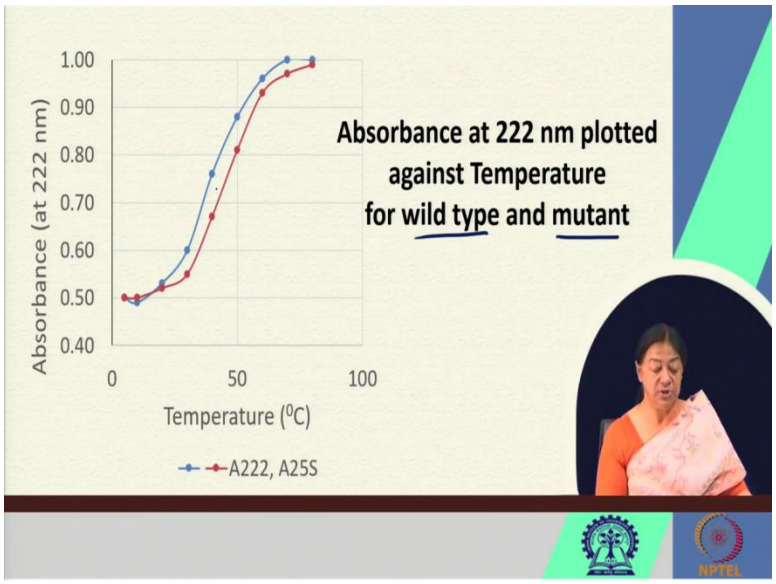
What do we observe, that at low temperatures when the protein is folded the absorbance is 0.5. This means that, this is the baseline measurement which gives the absorbance of the solution when the protein is in its folded form; that means, there is no part of the protein that is unfolded. Let us now plot the data and see what it looks like.

(Refer Slide Time: 20:51)



This [refer to slide] is the plot of the raw data. What do we have? Along the x axis we have the temperature, along the y axis we have the absorbance at 222 nm and we saw the initial value at 0.5 and this value at 1. So, this gives us an indication that with temperature the absorbance value follows this curve that has been drawn. This absorbance at 222 nm is plotted against the temperature for the wild type protein.

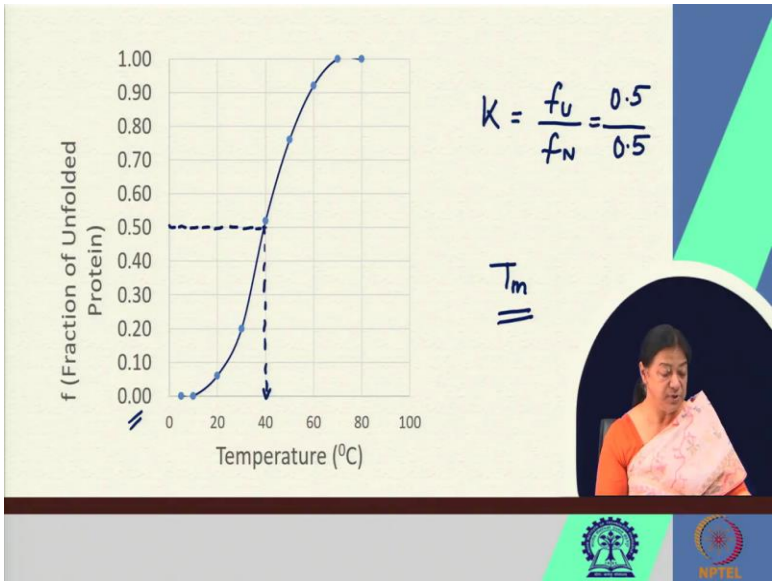
(Refer Slide Time: 21:29)



We can do the same for the mutant. We have the absorbance at 222 nm again versus the temperature, but now we have one for the wild type and we have one for the mutant. And what do we observe? We have the wild type and we have the mutant; we see that there is a shift in the value of the temperature, the absorbance versus temperature curve.

What does that mean? The shift towards a higher temperature means that this is more stable; it does not melt as easily and the shift is not much. But it indicates stability, so this is how we can determine the Tm or the melting temperature.

(Refer Slide Time: 22:22)



Now if we were to plot the fraction of the unfolded protein, what do we see? We see that initially there is no fraction that is unfolded at the lower temperature where the measurements have been taken at 5°C. Then the highest temperature that was measured was 80°C and at that point we

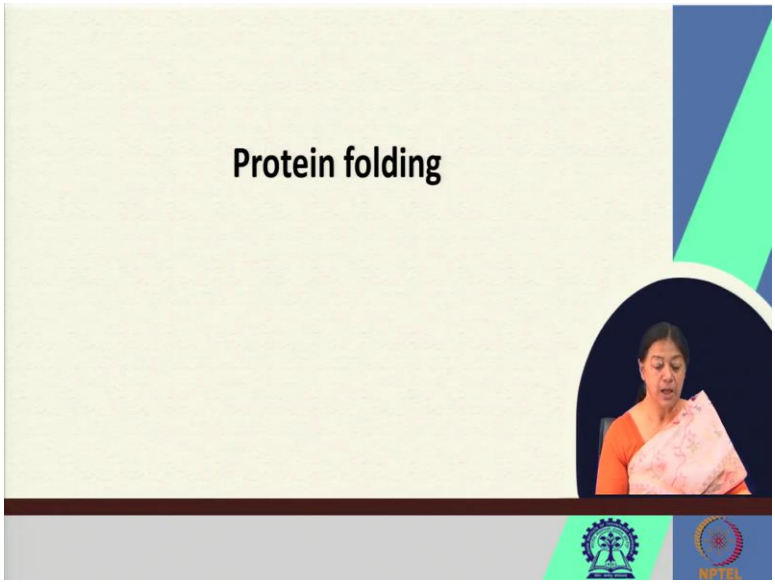
indicate that there is complete unfolding of the protein. When we go along this particular fraction, we are looking at a fraction unfolded, looking at a fraction in the native form.

If we look at the K value of this and say we have this at 0.5, indicating that 50% of it is folded and 50% of it is unfolded, we have a K value of 1, indicating a ΔG value of 0. So what do we have is, we can consider the 0.5 value that we have here for the fraction folded and from this we can find out that this is the melting temperature of the protein.

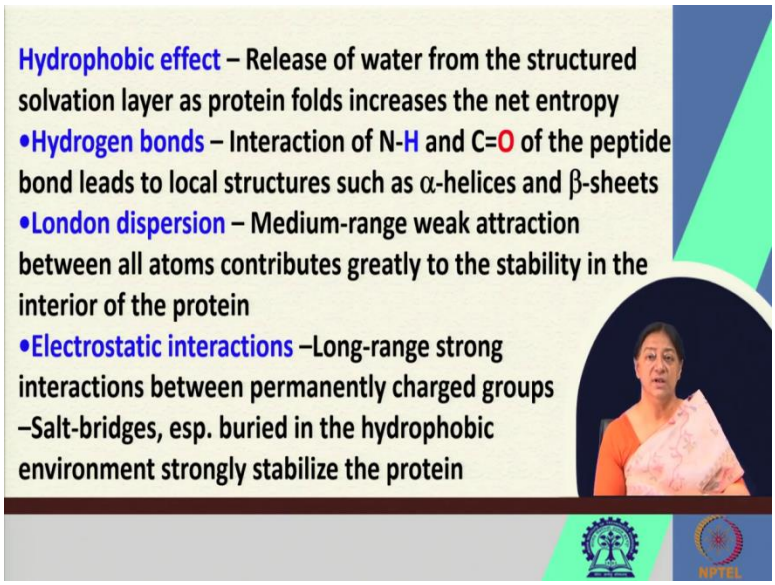
In this manner we can determine the Tm values of the proteins, that give us an idea of the stability of the protein. There are several ways in which we can measure the protein folding/unfolding. There are various techniques to measure these. We saw one where we looked at the variation in the fluorescence spectra and here we looked at a variation in the absorbance at a specific wavelength, which would give us the information about the unfolding of the protein.

We have to look at a specific value or a specific parameter that is going to give us an indication of the change in the protein structure, that is going to be related to the protein stability. We can also consider an enzyme and denature the protein by whatever method we consider. We can also look at the activity of the enzyme to understand the changes in the structure of the protein, because we know that for the enzymatic activity to occur we need the folded structure of the protein, because we want to have the active site in its perfect three dimensional conformation.

(Refer Slide Time: 25:00)



(Refer Slide Time: 25:04)



So in our protein folding understanding, we looked at the hydrophobic effect. That is the release of the water from the structured solvation layer as the protein folds, which increases the net entropy. The hydrogen bonds, that are the interaction of the N - H and the C = O of the peptide bond leading to specific local structures, such as α-helices and β-sheets.

The London dispersion forces, that were medium range weak attraction between all atoms that also contribute greatly to the stability in the interior of the protein. And the electrostatic interactions that were long range strong interactions between charged groups giving us these iron pairs and the salt bridges, especially those buried in the hydrophobic environment, strongly stabilize the protein.

These effects on protein structure and protein folding, that result in a stable three dimensional structure, which is the basis of protein chemistry and an understanding of all the interactions and all the functionalities that we are going to see in the subsequent lectures.

(Refer Slide Time: 26:27)

This is the end of module 3 protein structure and stability.

Thank you.