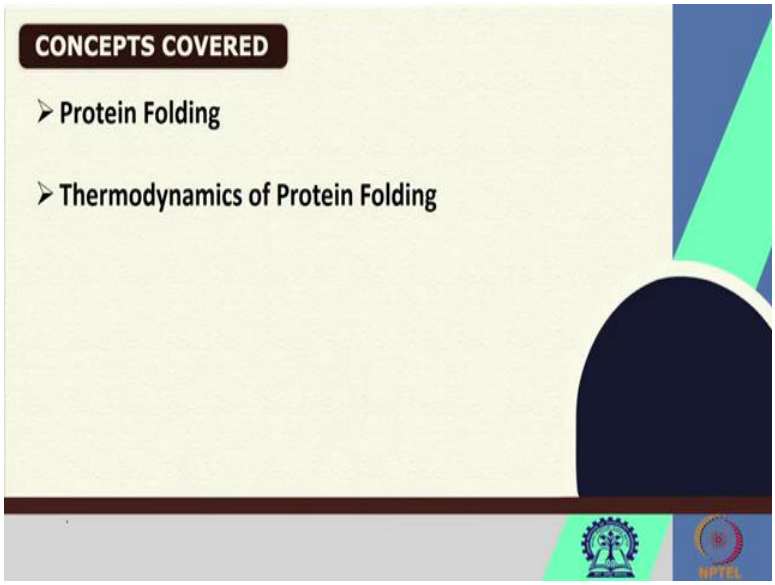
Fundamentals of Protein Chemistry
Prof. Swagata Dasgupta
Department of Chemistry
Indian Institute of Technology, Kharagpur

Module - 03
Protein Structure and Stability
Lecture - 12
Thermodynamics of Protein Folding

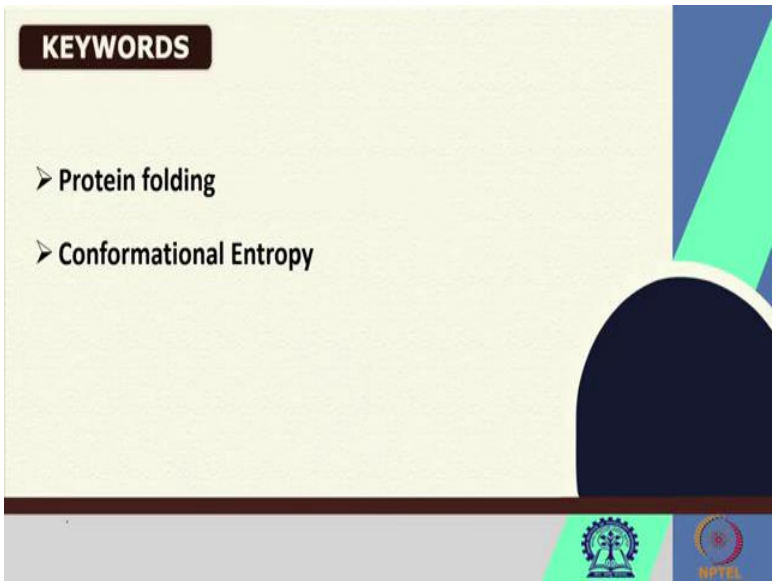
In the module on protein structure and stability, we will begin our discussion on the thermodynamics of protein folding. This is an important aspect and it encompasses a lot of the interactions and the studies that we have seen earlier.

(Refer Slide Time: 00:30)



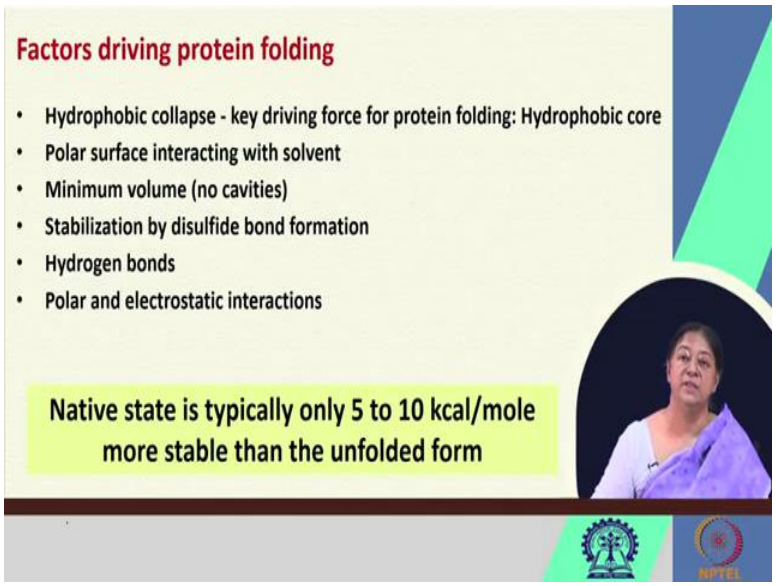
So if we look at the concepts that are going to be covered they are going to be protein folding in general and the thermodynamics of protein folding.

(Refer Slide Time: 00:37)



And this is going to have an important aspect in the conformational entropy that we will see as we go along.

(Refer Slide Time: 00:46)



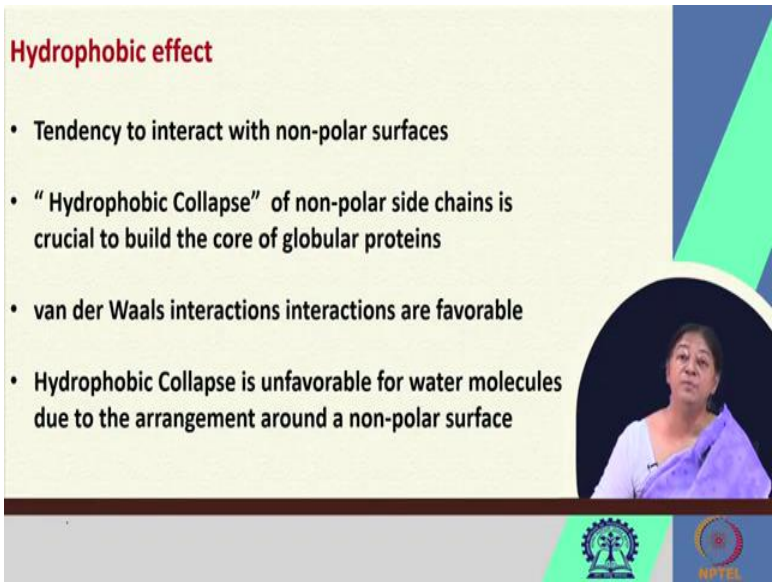
The factors that actually drive protein folding are hydrophobic collapse that results in the formation of the hydrophobic core, we have the polar surface that interacts with the solvent; we have a minimum volume that is occupied, so that it is a compact structure; then we have the stabilization by the other covalent bond formation that is the disulfide bonds.

Now the additional aspect is hydrogen bonds that can occur as we saw in the secondary structure elements between the C = O and the NH of the backbone. In addition, we can have these hydrogen bonds between the side chains and with the polar solvents.

Considering that this is an important aspect each of these are going to contribute thermodynamically to the aspects of the protein folding that we are going to see. And of course, we have the polar and the electrostatic interactions, usually between the ion pair interactions that we see between the acidic and the basic amino acid groups.

Now, the native state is typically only 5 to 10 kcal/mole, more stable than the unfolded form which we will be discussing.

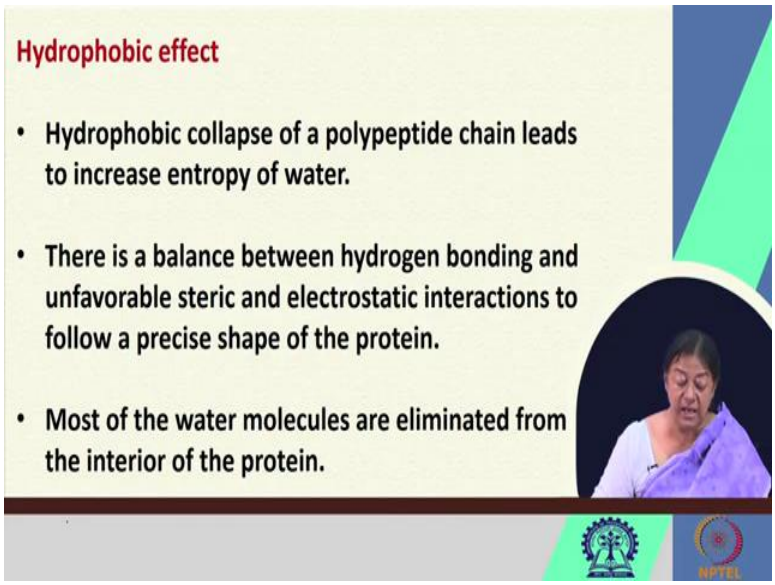
(Refer Slide Time: 02:04)



When we look at the hydrophobic effect, what exactly does it mean? It means that this is a tendency of non-polar surfaces to interact with each other to sequester themselves away from the solvent.

The hydrophobic collapse that we see of the non-polar side chain is crucial to build the core of the globular proteins, because the van der Waals interactions that we see are favorable. But nevertheless the hydrophobic collapse is unfavorable for water molecules, because they have to arrange themselves around the non-polar water surface, but we have to remember that we are looking at amides here.

(Refer Slide Time: 02:40)

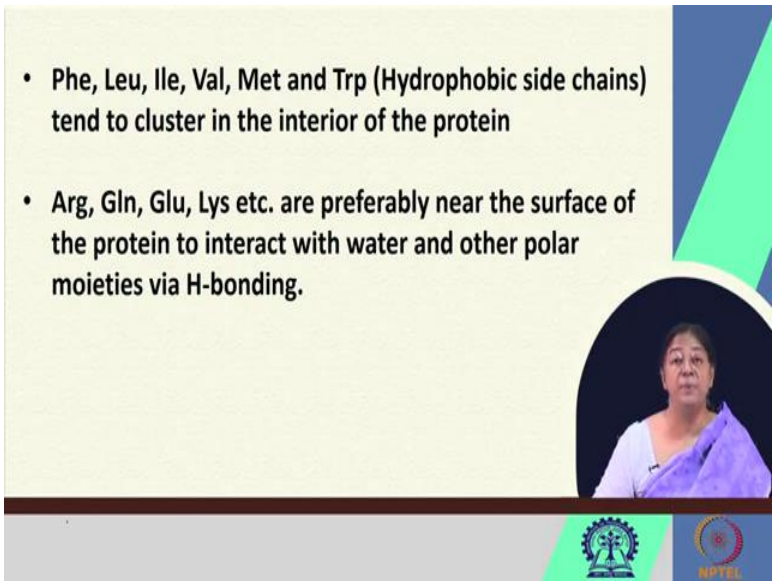


Each peptide bond is an amide bond that is linked by a C = O and an NH. So, this can actually interact with the water. When we look at the hydrophobic collapse of a polypeptide chain, we see an increased entropy of water; in the sense that the water that had been collectively around the polypeptide chain in its unfolded form, now has to rearrange itself to form the polar contacts with the molecules or rather with the residues that are on the surface of the protein. And, in addition to what happens to the core of the protein; there are these hydrophobic amino acid residues that tend to move away from the surface of the protein.

What happens is, there is a balance between the hydrogen bonding and the unfavorable steric and electrostatic interactions that follows the precise shape of the protein and most of the water molecules in this case are eliminated from the interior of the protein.

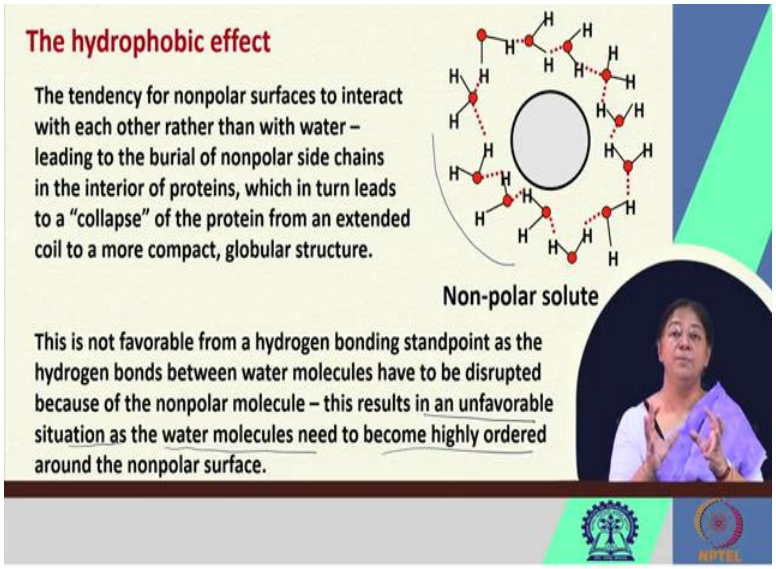
But nevertheless, there are some cases where we see, what are called structured water molecules that are hydrogen bonded in such a fashion that they are more or less rigid in their mobility, in terms of being a water molecule or with respect to the ones that we can see on the surface that are more flexible in nature.

(Refer Slide Time: 04:17)



When we look at the non-polar amino acid residues for example, phenylalanine, leucine, isoleucine, valine methionine and tryptophan to certain extent, we see they tend to cluster in the interior of the protein. And the other amino acids are preferably near the surface of the protein, so that they can interact with the water and other polar moieties by hydrogen bonding formation.

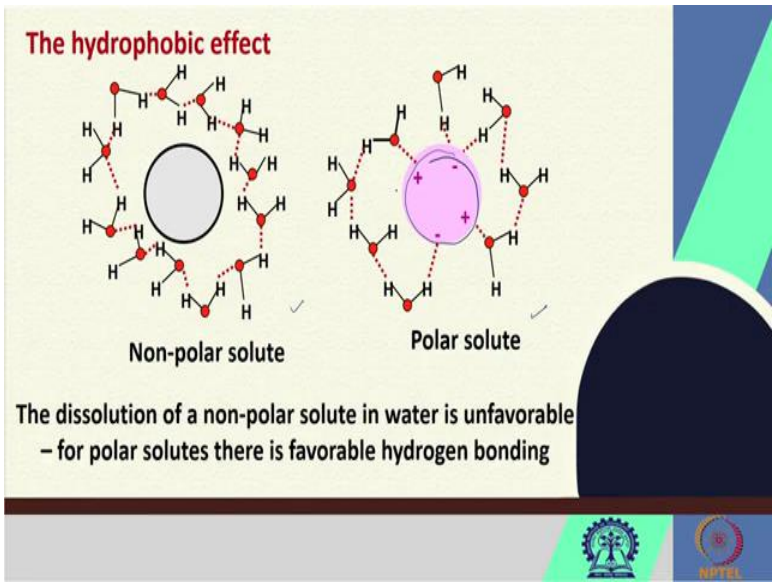
(Refer Slide Time: 04:43)



So if we look at the nonpolar solute now, we observe that there are hydrogen bonding network around this non-polar solute. There is a tendency for the non-polar surfaces to interact with each other rather than water, that is going to lead to the burial of the nonpolar side chains is what is termed as the collapse. So the hydrophobic side chains that are there in this unfolded protein will tend to come together in a collapsed form to form the core of the protein from an extended coil that did exist.

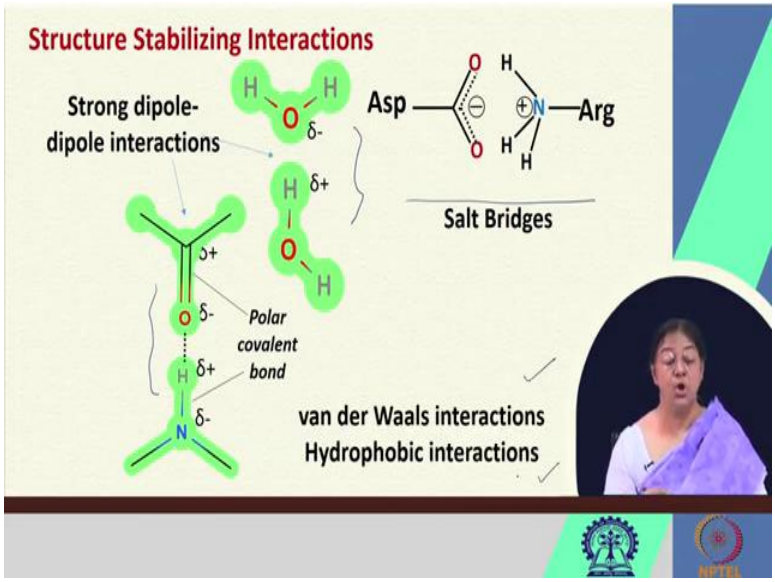
If we look at it from a hydrogen bonding standpoint, it is not actually very favorable. The hydrogen bonds between the water molecules have to be disrupted, because the non-polar molecule is there. And what happens is, an unfavorable situation results as the water molecules need to become highly ordered to bring about or to sequester themselves away. When the nonpolar parts sequesters itself away from the water, then the rest of the water molecules have to arrange in an ordered manner around it.

(Refer Slide Time: 06:08)



This gives an entropic aspect. When we look at a non-polar solute [refer to slide], this is what we see. When we look at a polar solute, we know that the polar molecule or the polar moiety itself has hydrogen bonding capability with the water molecules. So the dissolution of the non-polar solvent in water is unfavorable and for polar solutes, we have favorable hydrogen bonding.

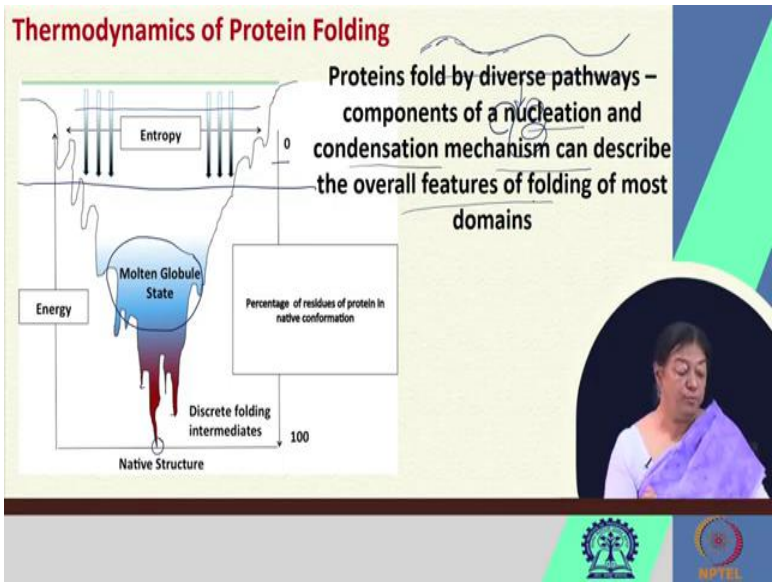
(Refer Slide Time: 06:35)



When we look at the other structure stabilizing interactions, we see that there are strong dipole-dipole interactions, which we would expect with the water molecules and of course again with the hydrogen bonding, that we see in the C = O, N-H of the peptide linkages throughout the protein. So, whether we are looking at a β-sheet or the strands of the β-sheet connecting together or in the α-helix, we see this sort of interaction between them.

The other salt bridges that are observed are between the acidic and the basic amino acid side chains. In addition to this we see van der Waals interactions and hydrophobic interactions. All of these are structure stabilizing interactions.

(Refer Slide Time: 07:22)



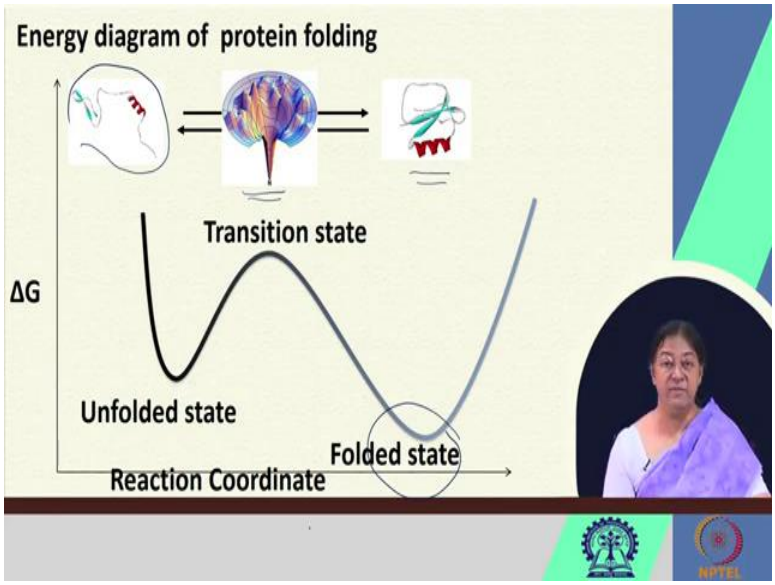
Now if we look at the thermodynamics of the protein folding, we realize that when we have the unfolded form that is completely in a random state, then what is going to happen. There are going to be many many such conformations possible.

Each of these have the peptide bond that link the different amino acids together. So, they have a water network around it, but then when we have the hydrophobic parts come together in the collapse, this water is disrupted and we have the entropy that is going to be large here [refer to slide], gradually decreasing in terms of looking it up from a protein standpoint.

So, this protein molecule is now going to fold into a compact form and it is going to form a native structure. This native structure gives different folding intermediates in between a molten globule state that says there are some elements or some nucleation that occurs, where we have some β strands form to form β-sheets; a part of a helix and so on and so forth. And then we have the compact native structure.

Initially when we have the native conformation, we have the polypeptide chain, we have a large number of possibilities that gradually then get to a minimum energy state to give us the native structure. So when we look at the protein folding, it occurs by diverse pathways and there are components of a nucleation and a condensation mechanism, that actually describes the overall features of folding of most of the domains that we know about.

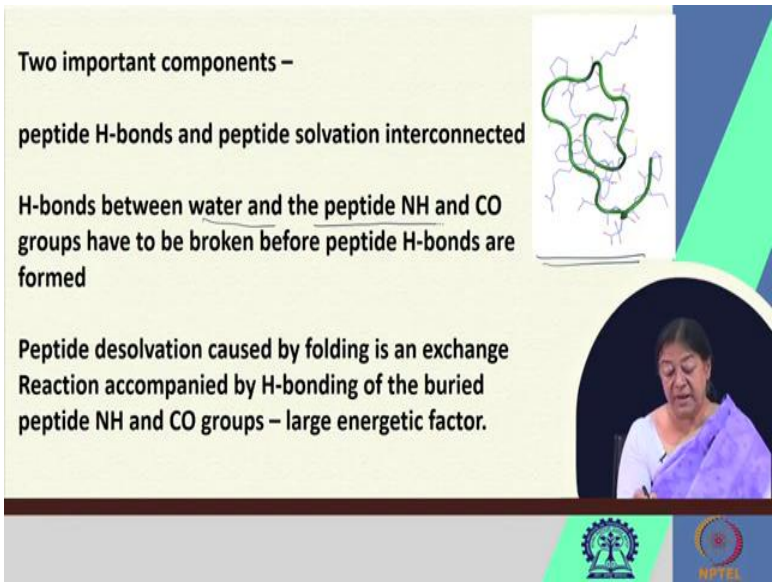
(Refer Slide Time: 09:18)



If we look at the unfolded structure in this form [refer to slide] we look at a protein folding funnel and we have the final folded structure. So the unfolded structure and the folded state are in a energy form, where we have the folded state at a minimum energy form that is going to be the compact structure.

But there is much more to the protein folding aspects or the thermodynamics of the protein folding that will be covered in this course.

(Refer Slide Time: 09:51)

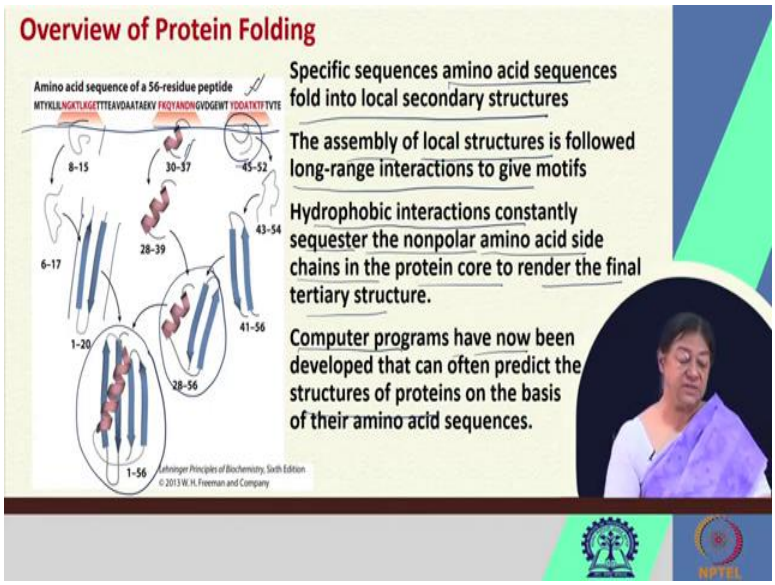


There are two important components, the peptide hydrogen bonds and the peptide solvation that are interconnected.

So, the hydrogen bonds that we saw between the water and the peptide NH and CO groups, is the unfolded form that is on the right here [refer to slide], they have to be broken before the peptide H-bonds themselves are formed. The peptide desolvation caused by the folding, is an exchange reaction that is accompanied by H-bonding, hydrogen bonding of the buried peptide NH and CO groups, that results in a large energetic factor.

This is important in the peptide solvation and the hydrogen bond formation that we are talking about.

(Refer Slide Time: 10:49)



If we look at the overview of protein folding. We look at an amino acid sequence of a 56 residue peptide and we see the steps that are involved. So the specific sequences, the amino acid sequences fold into local secondary structures.

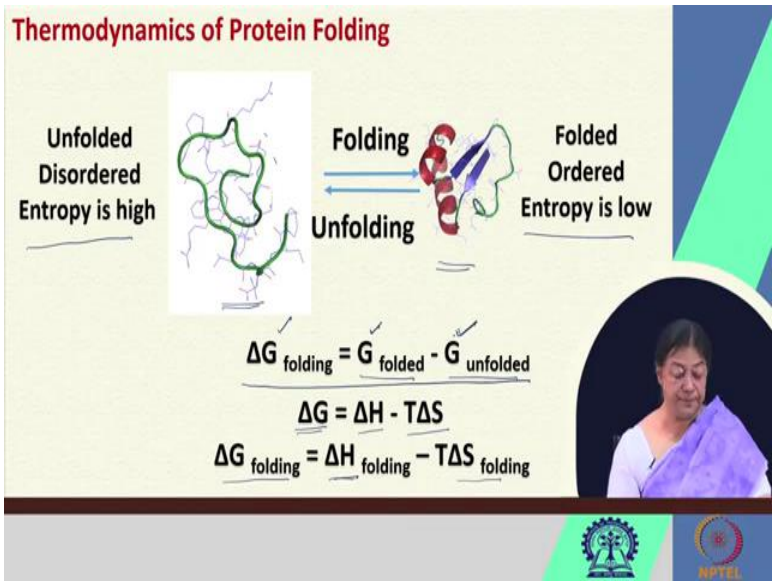
What do we mean by this? We have certain regions that are marked in red here [refer to slide] and we see that this part has formed a coil or a turn like this, this part is a helix and then this part gradually gets into strand formation to form a β-sheet. Then what happens is, there is the assembly of the local structures, that is followed by the long range interactions to give specific types of motifs that we looked at.

For example, we can have a βαβ motif or an αβα motif. So that would be again a secondary structure element coming together, where we would have a β-sheet and we would also have an α-helix. While this is happening, the hydrophobic interactions constantly sequester the non-polar amino acid side chains in the protein core to give its final tertiary structure.

So, while the hydrogen bonds are being formed, the water molecules are being pushed out from the central core of the protein, so that the hydrophobic interactions may take place. Then that is going to give rise to the protein core. Now we have computer programs that have been developed, that can often predict the structures of proteins on the basis of this information.

On the basis of the amino acid sequences, we might be able to say that given this sequence because we know from Anfinsen's experiment, that all the information that is required for the three-dimensional structure of the protein is present in this amino acid sequence.

(Refer Slide Time: 12:54)



So if we look at the thermodynamics of the folding, now we have our unfolded structure. We remember again that some of these on the surface are still hydrophobic in nature.

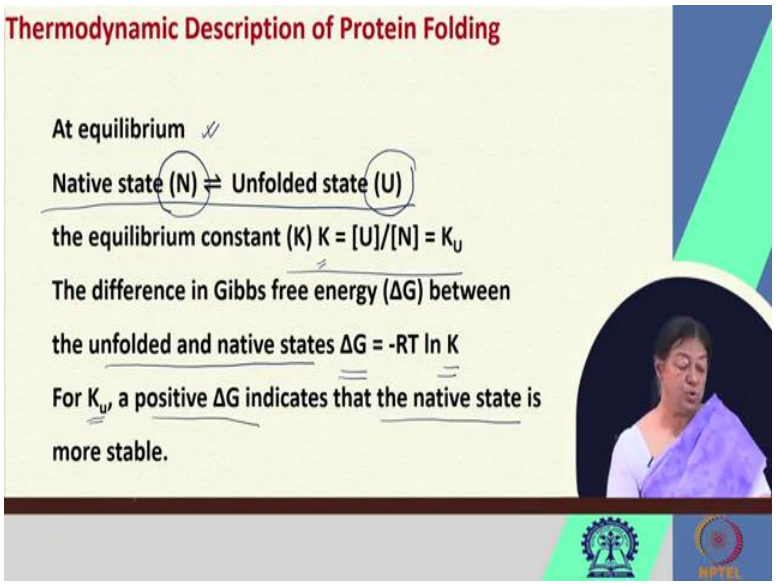
But it gradually gets to a nucleation and a condensation, giving us the folded structure. If we want to look at the free energy of folding, it would be the final form; that is the G_{folded} - $G_{unfolded}$. So, we look at the unfolded formation and we see that it is disordered and the entropy is high. On the other hand, if we look at the folded conformation it is an ordered conformation.

The entropy now is low. If we look at the formation, we know that the free energy ΔG is given by ΔH - $T\Delta S$. So, if we look at the $\Delta G_{folding}$, we have to look at a $\Delta H_{folding}$ and a $T\Delta S_{folding}$. So, if we now look at what the $\Delta S_{folding}$ is going to be in terms of what we see here? It is going to be;

$\Delta S_{folding} = S_{folded} - S_{unfolded}$ and

$\Delta H_{folding} = H_{folded} - H_{unfolded}$.

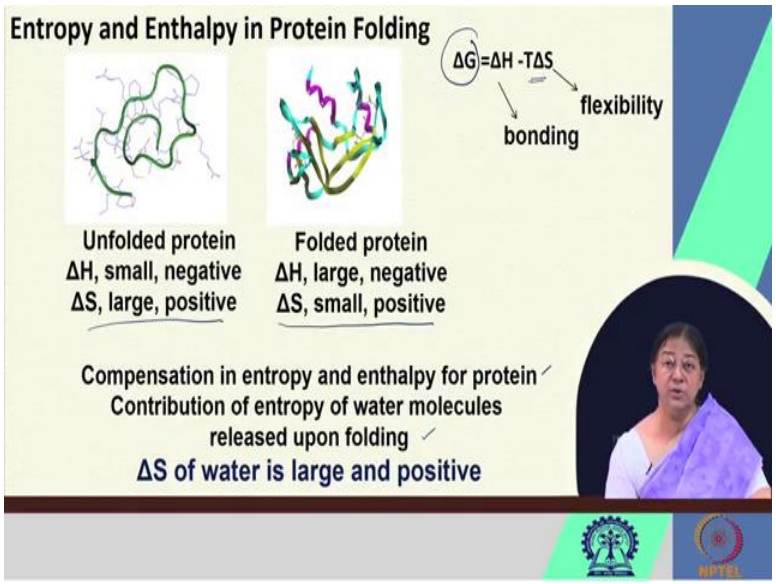
(Refer Slide Time: 14:22)



At equilibrium the native state and the unfolded state are reversal. At equilibrium the equilibrium constant (K) K = [U]/[N] = $K_{U,}$ which you will see when we look at protein denaturation as well.

So, the difference between this gives us the $\Delta G_{unfolded}$ and the native states, given that this is in equilibrium, we have an equilibrium constant associated with this and we have the ΔG that we can calculate. From this we can actually get the ΔG of the folding process. When we have a $K_{unfolded}$, we have a positive ΔG that indicates that the native state is more stable.

(Refer Slide Time: 15:20)



Now when we look at the protein folding process, we say it is a spontaneous process, this means that the ΔG value has to be negative. If it is negative and we know that from a protein standpoint, the flexibility that we observed in terms of ΔS, is taking it from a more disordered state to a more ordered state, giving us a ΔS that is actually negative.

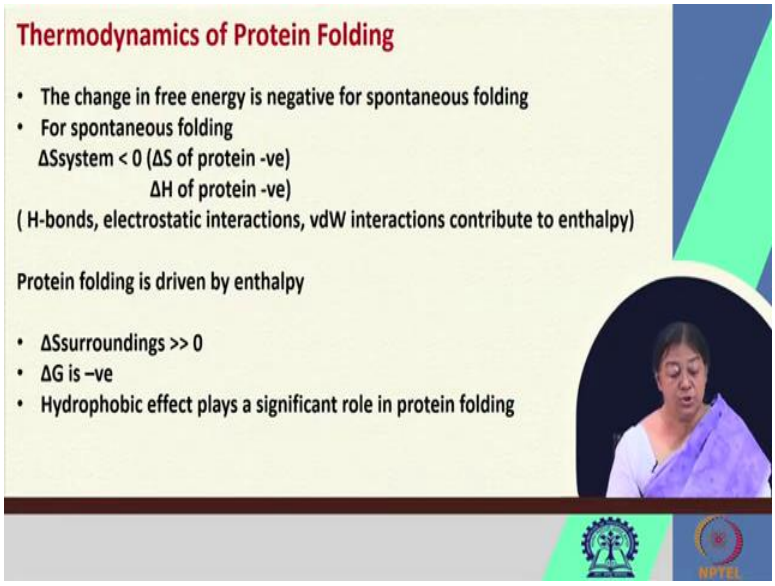
What happens in this case? When we are looking at the -TΔS form, it is actually a positive quantity. When we look at the bonding, the enthalpy contributions for the different bonds that are possible to form, they give us our contributions to the overall protein folding.

So again for the unfolded protein, we have a ΔH that is small and negative and we have a ΔS that is large and positive. When we now look at the folded protein, we have the ΔH that is large and negative and a ΔS that is small and positive. Now for the folded part we realize that the ΔS is small and positive and it is going to be large and positive, because we are disrupting the water molecules.

There is a conformational entropic contribution there. For the compensation in entropy and enthalpy for the protein, there is a contribution of entry of the water molecules that is released upon the folding. So, the ΔS of the water that we actually see of the surroundings is large and positive.

All of this together is going to contribute to the overall ΔG that we observe.

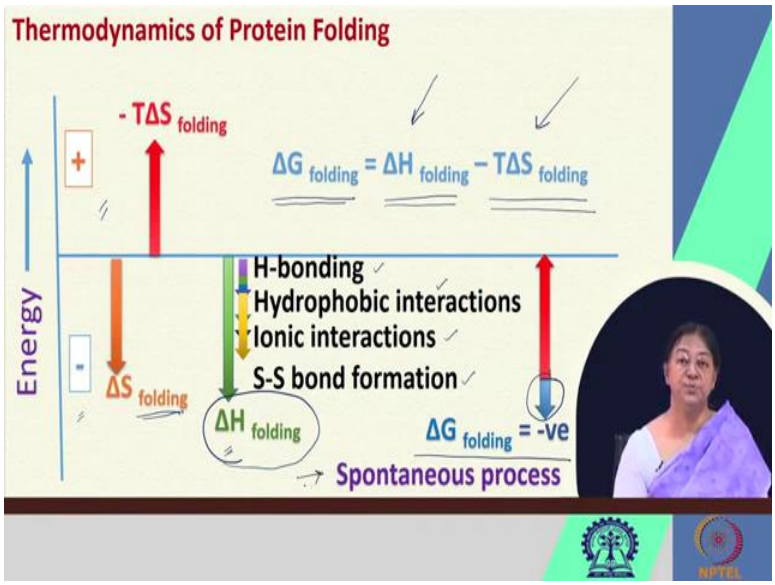
(Refer Slide Time: 17:11)



If we look at the change in the ΔG, the change in the free energy, it is going to be negative for the spontaneous folding. So for the spontaneous folding, what we are going to observe is we have the ΔS of the system, which is our protein that is going to be negative; Why? Because we have the disordered state that folded into an ordered manner, giving us a ΔS that is negative.

The ΔH of the protein is also negative because, we have all the enthalpy contributions from all the possible interactions; the hydrogen bonds, the electrostatic interactions, the van der Waals interactions, they all contribute to the enthalpy. The protein folding that we have is driven by enthalpy, but the ΔS of the surroundings is much greater than 0; the ΔG is negative and the hydrophobic effect plays a very important role in the protein folding process.

(Refer Slide Time: 18:13)



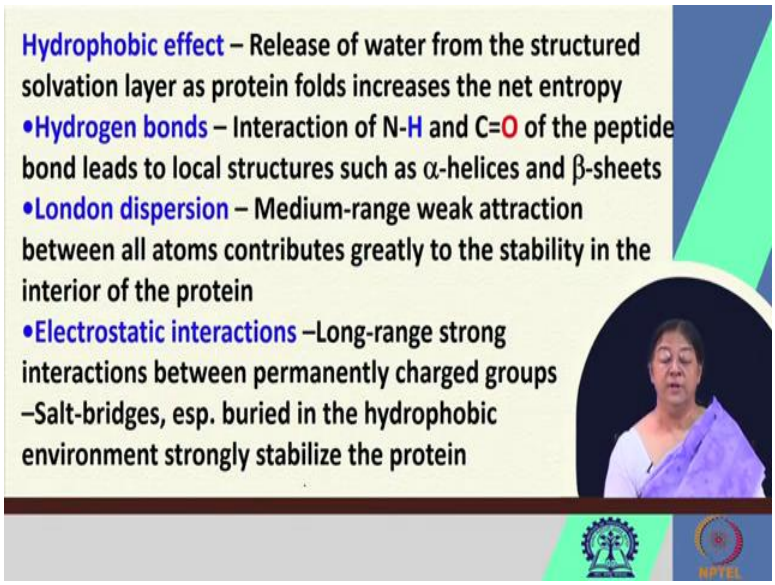
So if we look at the contributions to $\Delta G_{folding} = \Delta H_{folding} - T\Delta S_{folding}$. And we found out, that if you want to look at the $\Delta H_{folding}$, we have to look at $H_{folded} - H_{unfolded}$. When we look at the $\Delta S_{folding}$, we have to look at the $S_{folded} - S_{unfolded}$. Because when we look at folding our final form is the folded form.

If we look at the contributions now, when we are look at $\Delta S_{folding}$, we see that this contribution is negative, meaning that we have a negative contribution to the ΔS, because we are going from a disorder form to an ordered from. So when we look at the expression in the $-T\Delta S$, this becomes a positive quantity in the $-T\Delta S$ contribution.

However, when we are looking at the $\Delta H_{folding}$, that is the enthalpy contributions to protein folding, there are all contributions that are going to add up. What are these contributions? The contributions correspond to the hydrogen bonding that are possible between the polar entities; the hydrophobic interactions that are possible between the non-polar entities that are going to form the hydrophobic core of the protein; we have the ionic interactions that are possible; we also have the disulfide bond formations that are possible between the cysteine residues.

And what is going to happen here, is all these are going to contribute to the overall enthalpy of the folding that we see. In effect when we look at the $\Delta G_{folding}$ and we look at the $\Delta H_{folding} - T\Delta S_{folding}$, we see that the enthalpic contributions give the small negative factor for the $\Delta G_{folding}$, that makes $\Delta G_{folding}$ process be negative, which means that this is a spontaneous process.

(Refer Slide Time: 20:29)



So the contributions that we see are extremely important. If we can summarize what we have learned, we have the hydrophobic effect. This hydrophobic effect is the release of water from the structured solvation layer, as the protein folds increases the net entropy. So we have to be very careful here, as to why are we considering an increase in the net entropy.

The water molecules that are being moved away from the solvation of the peptide are now getting folded because we have the hydrophobic interactions occurring between the non polar residues that would tend to remain away from the water; tend to remain away from the solvent. So this is what the hydrophobic effect is.

So it is the release of the water from the structured solvation layer as the protein folds, that results in an increase in the net entropy. But from a protein standpoint, when we are looking at the disordered polypeptide chain, when that is getting folded into a more ordered compact structure, that is resulting in a decrease in entropy for the system-that is our protein chain. When, we look at the hydrogen bonds, we look at the interaction of the NH and the CO of the peptide bonds, leading to the local structures such as α-helices and β-sheets.

So in the overview of the protein folding we saw the formation of the α-helices, the formation of the β strands that form the β-sheets; that result in the local structures. And then we have the formation of the domains that are then going to lead to what we call the molten globule state. This molten global state is going to gradually get to an energy minimized form, where we have the native state being the most stable state that we know as the folded state of the protein.

In the London dispersion forces, we have medium range weak attraction between all atoms. And this also contributes greatly to the stability in the interior of the protein. We have to remember that this is the dipole-dipole interactions. So there is an instantaneous dipole that occurs in all atoms, which brings about van der Waals type of interactions.

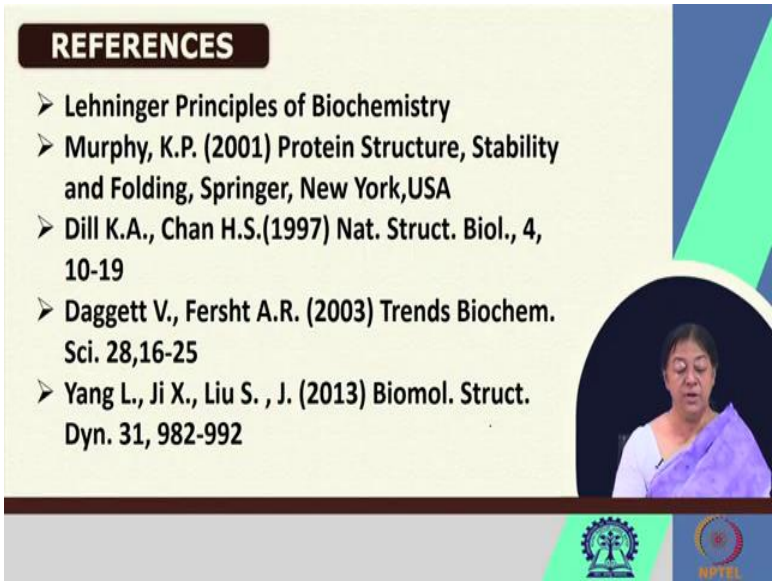
We have the electrostatic interactions. Now, the electrostatic interactions are the long-range strong interactions between permanently charged groups. Which are the permanently charged groups? These are the ones that have the positive charge or the negative charge. So we have the

basic amino acid side chains, we have the acidic amino acid side chains. And the interactions between these, are strong interactions and these ion pairs that are formed, the salt bridges especially buried in the hydrophobic environment, strongly stabilize the protein.

So when we have a specific type of amino acid residue that belongs to the charged polar type, that is the basic or the acidic type, they would prefer to be on the surface of the protein. But if it so happens that they are in the hydrophobic environment of the protein in the hydrophobic core, there is usually a salt bridge occurring where we have an interaction between the positive amino acid and the negative amino acid in the formation of the salt bridge.

Again, we can look at the water molecules that are present within this, where sometimes we have bridged water molecules that have a specific structured characteristic to them, because they are involved in hydrogen bonding and they are caged within the protein structure. But those on the surface are more flexible because they can exchange themselves with the bulk water.

(Refer Slide Time: 25:08)



So, the different references that we have followed are these [refer to slide] specific books and in addition, we will revisit protein denaturation to see what structural aspects occur, when the protein loses its native structure.

Thank you.