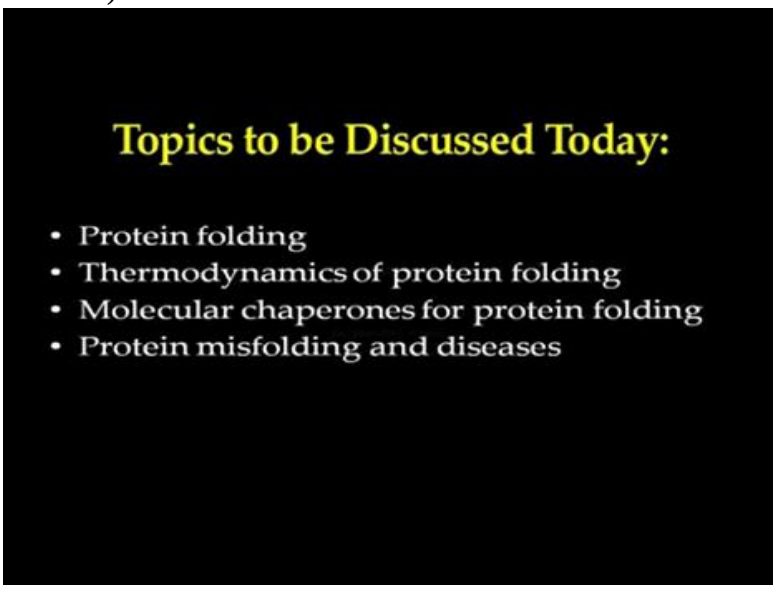
Proteins and Gel-Based Proteomics
Professor Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay
Mod 02 Lecture Number 4

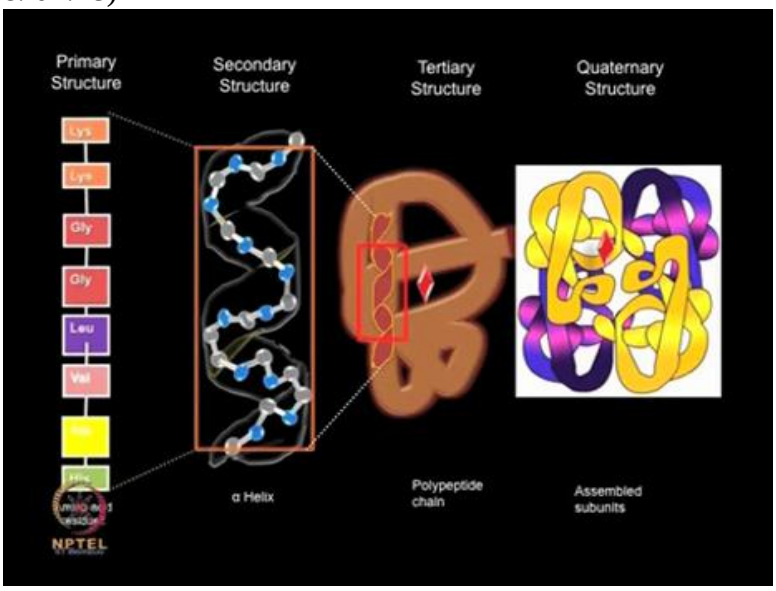
(Refer Slide Time: 00:16)



Today we will talk about proteins, its folding and misfolding. Understanding the processes of protein folding and misfolding has been a major research area from last several decades in biology, chemistry and physics.

Understanding protein folding and misfolding remains challenging and continues to motivate researchers to work both, experimentally and theoretically in this area. In today's lecture I will present and discuss the basics of protein folding, how this process works and how misfolding may result into various manifestations of diseases.

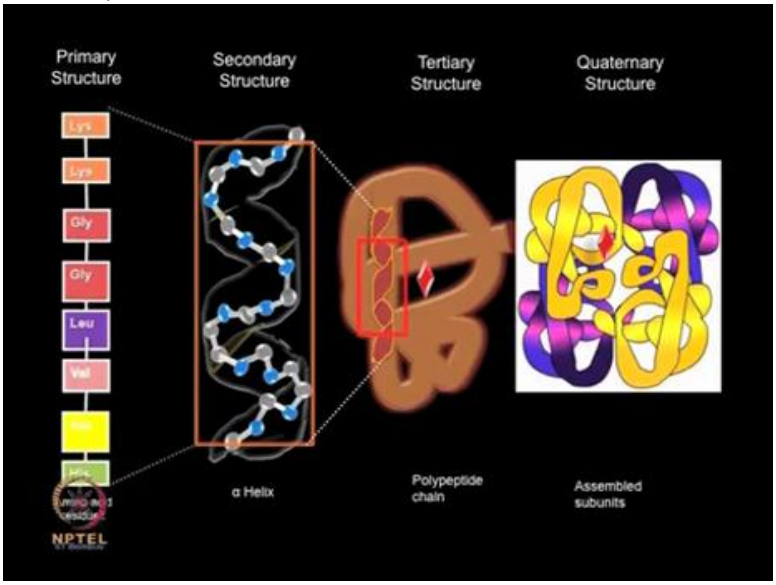
(Refer Slide Time: 01:18)



As we talked in the last class, proteins play a very crucial role in essential characteristics of living system. How they function, how they replicate themselves through the intricate molecular interactions. Proteins are most important classes of molecules which are involved in promoting and regulating essentially all the reactions which takes place in living systems.

We discussed previously the globular proteins. They can fold into conformations of ordered secondary and tertiary structures. The interactions which govern the formation of secondary, tertiary and quaternary structures involve different forces and interactions.

(Refer Slide Time: 02:24)

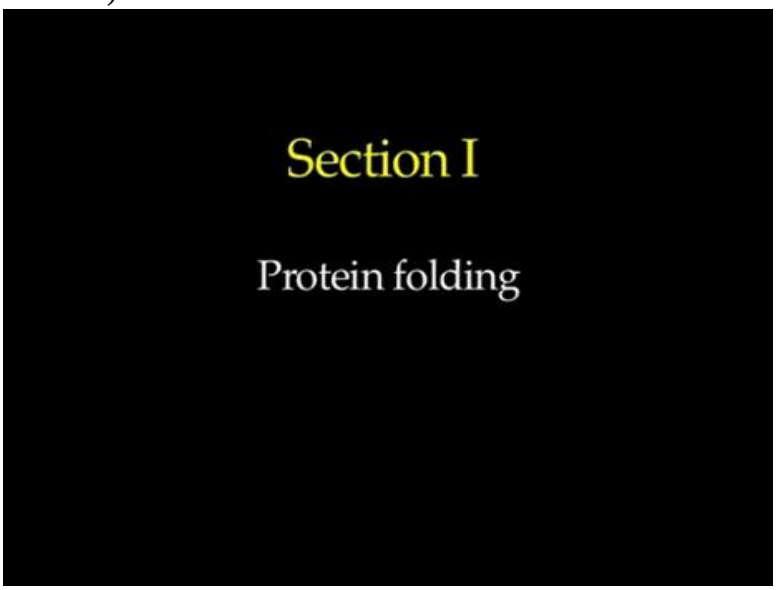


The cumulative effect of all of these interactions and forces are such that the folded proteins, the magnitude of the favorable reactions or interactions will be outweighing the sum of unfavorable interactions and as a result, it governs the protein folding.

(Refer Slide Time: 02:47)

So in today's lecture, we will talk about how protein folding occurs, some of the basic thermodynamics of protein folding concepts, how molecular chaperones govern the protein folding process, and how protein misfolding may result into various diseases.

(Refer Slide Time: 03:12)

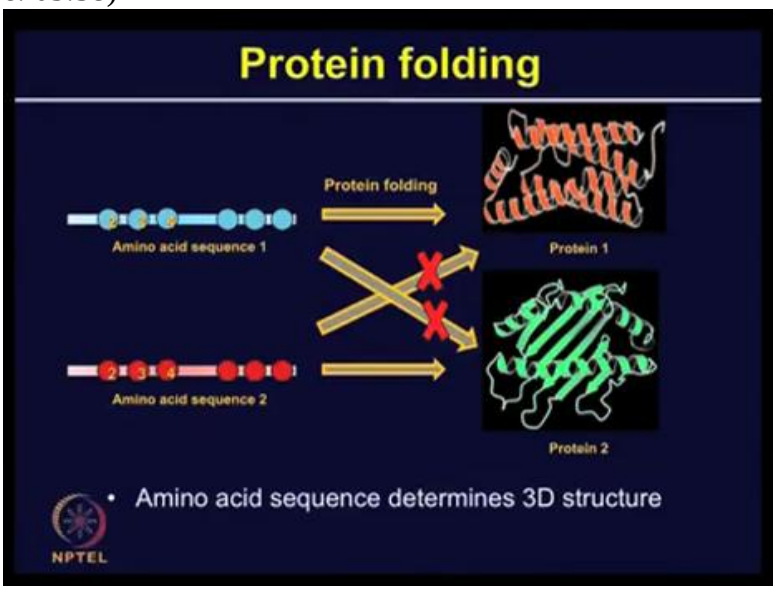


(Refer Slide Time: 03:18)

Let's talk about protein folding. Understanding the mechanism by which protein folding takes place still remains challenging for the scientific community.

Protein folding provides an elegant example of biological self-assembly and understanding such complex machinery provides very critical information not only for the understanding of protein folding but also the evolutionary aspects of proteins and various biomolecules.

(Refer Slide Time: 03:58)

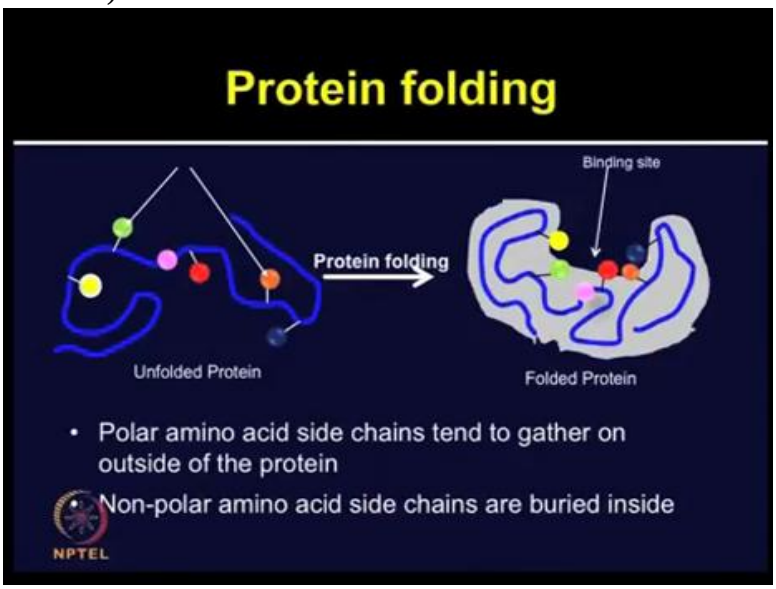


In protein folding, the amino acid sequence determines the three-dimensional structure. Now as you can see here, the proteins having very much specificity; if you have amino acid sequence 1 shown in blue color that will form protein 1 shown in the right side. If you have amino acid sequence 2 shown in red that will form protein 2.

Now if you take the amino acid sequence 1, protein 2 cannot be generated. Similarly if you take amino acid sequence 2, protein 1 cannot be generated. So there is very high specificity of amino acid sequence which can determine the three dimension structure of proteins.

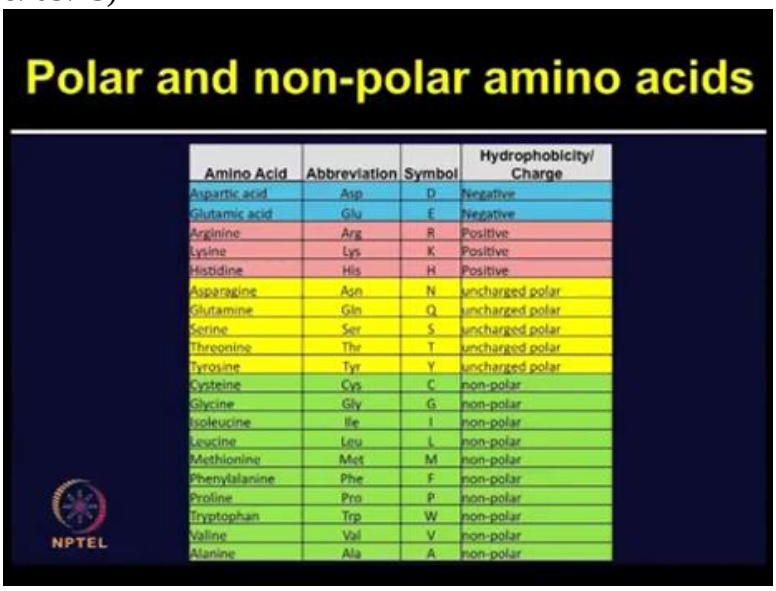
The protein folding process is governed by distribution of polar and non-polar amino acids. If you remember last class, we have talked about various amino acids.

(Refer Slide Time: 05:12)



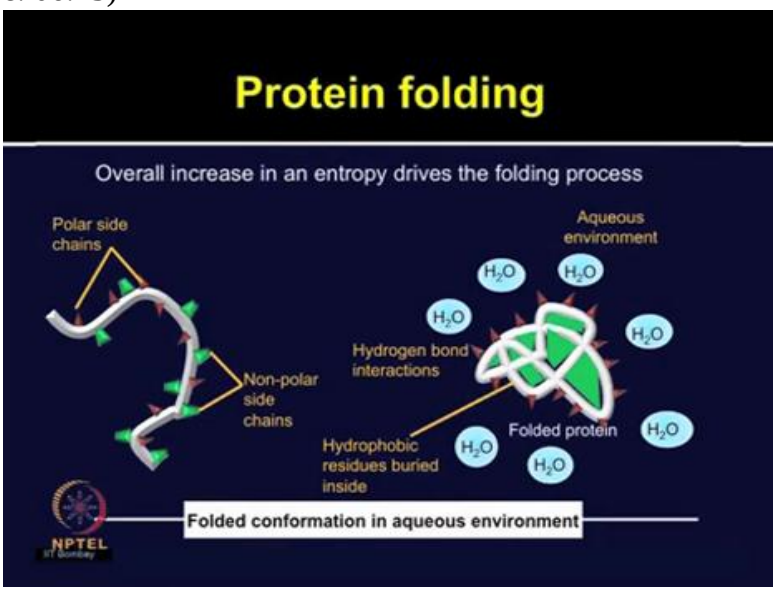
The polar side chains, they tend to arrange themselves near outside of the molecules. You take for example Arginine, Glutamine, Histidine. Similarly on the non-polar side chains they have tendency to cluster in the interior of molecules for example Phenylaniline, Leucine, Valine and Tryptophan.

(Refer Slide Time: 05:45)



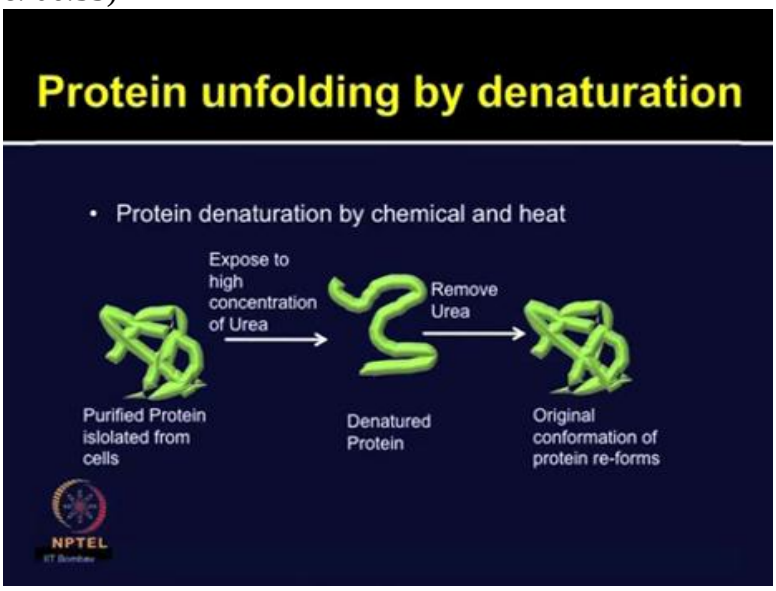
This chart is only for your information which shows there are various amino acids which belong to polar and non-polar category and then you can think of how they are going to govern the protein folding process. So continuing on protein folding, the hydrophobic amino acids; they are driven to associate the hydrophobic collapse.

(Refer Slide Time: 06:13)



So when these amino acids come together as you can see on the right hand side, the loss of water surrounding these amino acids increases entropy of the system. Therefore overall increase in entropy drives the folding process.

(Refer Slide Time: 06:35)

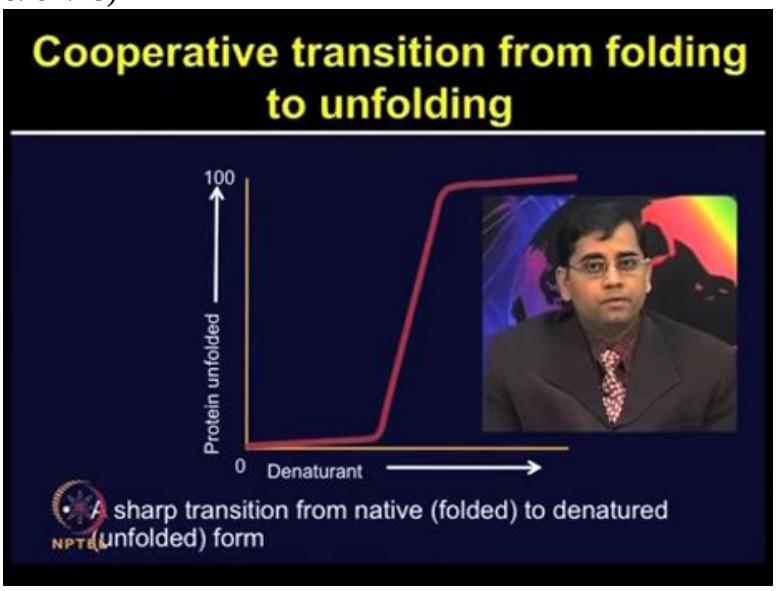


Now as we have seen in the classical experiment of Anfinsen, the protein unfolding can be done by using denaturants. So if you take denaturants, whether it is chemical like urea and Guanidium Chloride or you heat-treat it.

So as you can see here, if you have a purified protein isolate taken from the cells and you expose it to the high concentration of denaturants, whether it is chemical or heat, that will result into the denatured protein shown in the center.

If you remove the denaturing condition, it will again form the proper folding. Protein confirmation will be restored in its original form.

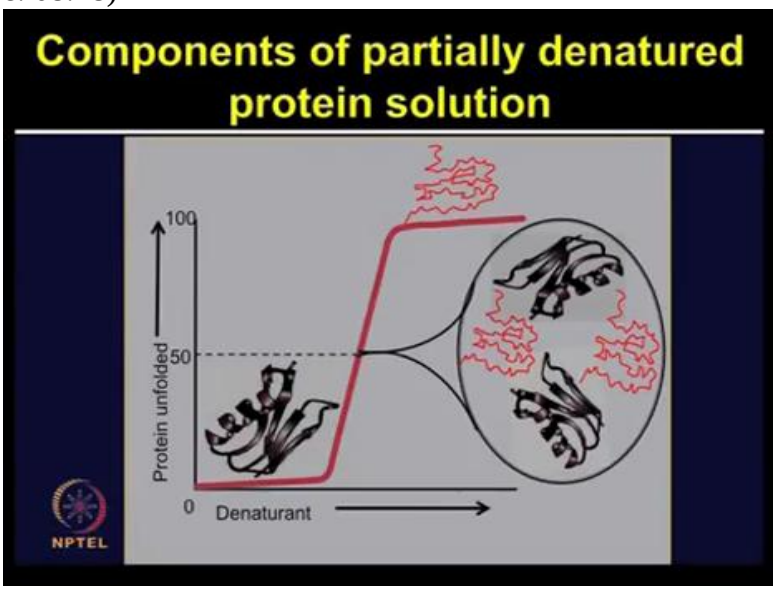
(Refer Slide Time: 07:28)



So how does this process of folding to unfolding works? There are various hypotheses and mechanisms have been proposed. Let's talk about cooperative transition here from folding to the unfolding form… as you can see in this graph, on the y--axis, the protein in the unfolded form from 0 to 100, and on the x-axis, the presence of denaturant.

A sharp transition from the native or the folded to the denatured or unfolded forms occurs. So only two conformational states are present significantly whether it is folded form or unfolded form. If denaturants are removed from the unfolded protein, it allows protein to make folded forms.

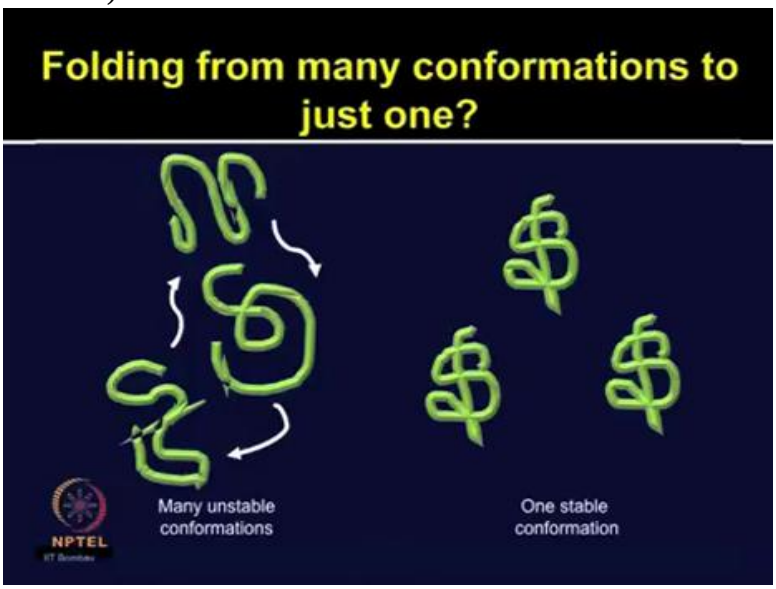
(Refer Slide Time: 08:28)



So what are the components of partially denatured protein? If you look at this graph, in the transition state at 50%, it will be 50% fully folded and 50% unfolded form of the protein.

However existence of only 2 states, the folded and unfolded or possibility of unstable transient intermediates between the folded and unfolded states still remains the topic of research in protein folding area. So how folding occurs from many conformations to only one form? The particular sequences along the polypeptide backbone, they impose key restrictions.

(Refer Slide Time: 09:30)

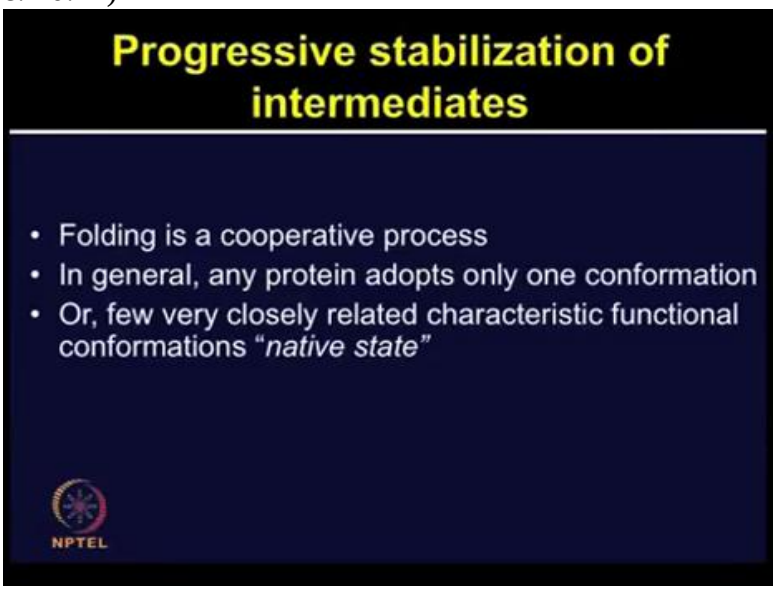


The various properties of the side chains which we have talked in the previous lecture including size, hydrophobicity, ability to form hydrogen and ionic bond, all of these govern this process.

Let's take example of Arginine. A side chain with positive charge might attract a segment of the polypeptide which has complimentary negative charge. You take, for example, aspartic acid. So these types of side chains and various types of backbone properties are going to impose key restrictions.

Therefore various types of folded conformations will be selected and it can result only the one which is going to govern the folding process.

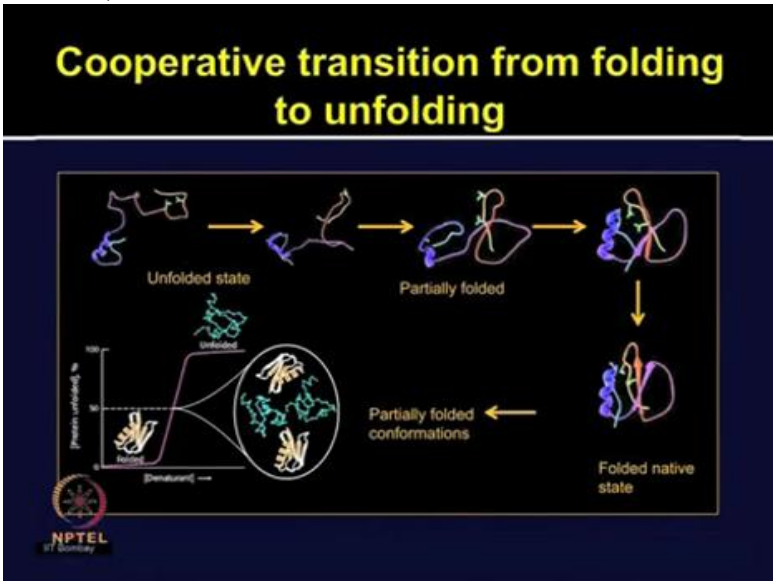
(Refer Slide Time: 10:22)



So there are various progressive stabilization of intermediates occur in the folding process. As we talked, folding is a cooperative process which involves progressive stabilization of various intermediates. In general any protein adopts only one conformation which we just talked in the last slide.

Or few very closely related characteristic functional conformations may occur which will give rise to the native state. Native state in this context here will be the conformation which has the lowest free entropy or the most stable folded form for majority of the proteins.

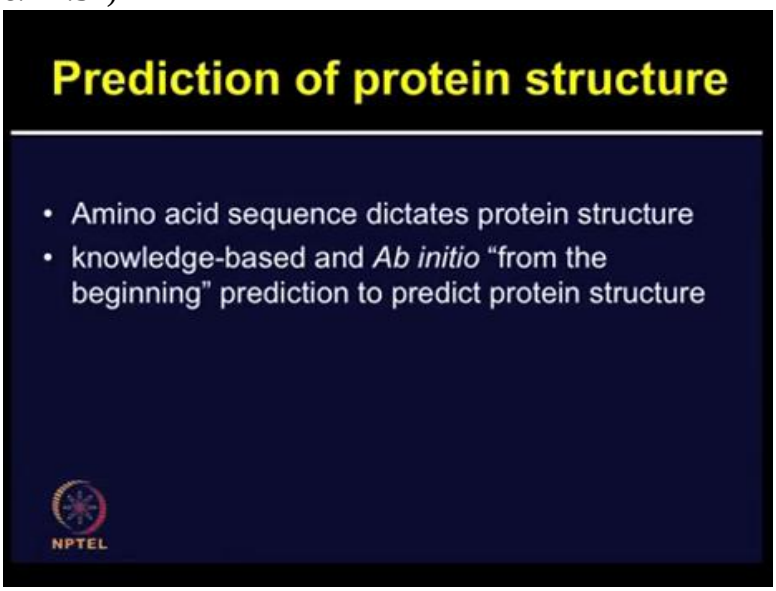
(Refer Slide Time: 11:16)



So let's see in the slide how cooperative transition occurs from the folding to unfolding. Folding is a cooperative process which arises from simultaneous formation of multiple

interactions within a polypeptide chain. If you take individually, each interaction is weak but their cooperative formation drives polypeptide chains towards the folded state.

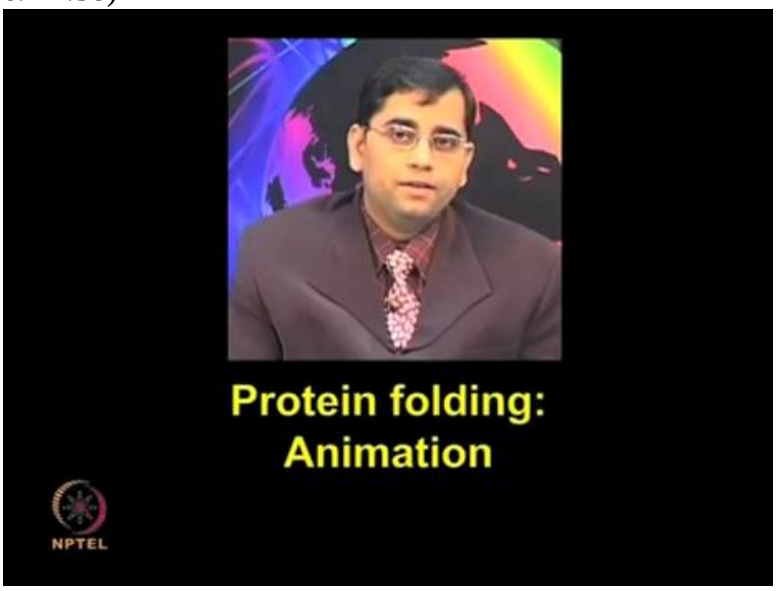
(Refer Slide Time: 11:51)



So, how to do structural prediction of proteins? As we have seen in the previous experiment, the amino acid sequence dictates the protein structure. So theoretically the prediction of final folded structure is possible from its sequence.

However there are long range of interactions and vast number of possible conformations which are possible and therefore it limits these types of predictions. However knowledge based and Ab-initio, from-the-beginning prediction do take place to predict the protein structure.

(Refer Slide Time: 12:38)

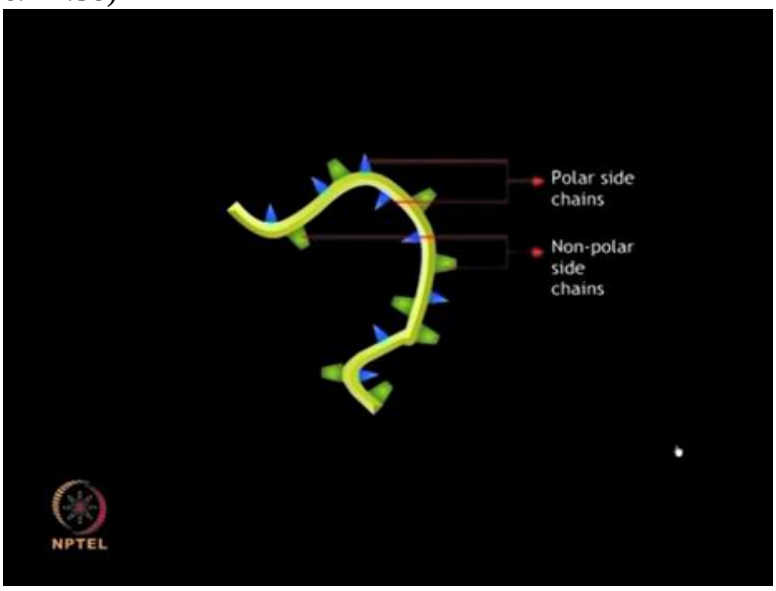


So let me show you the protein folding process in following animation.

(Refer Slide Time: 12:43)

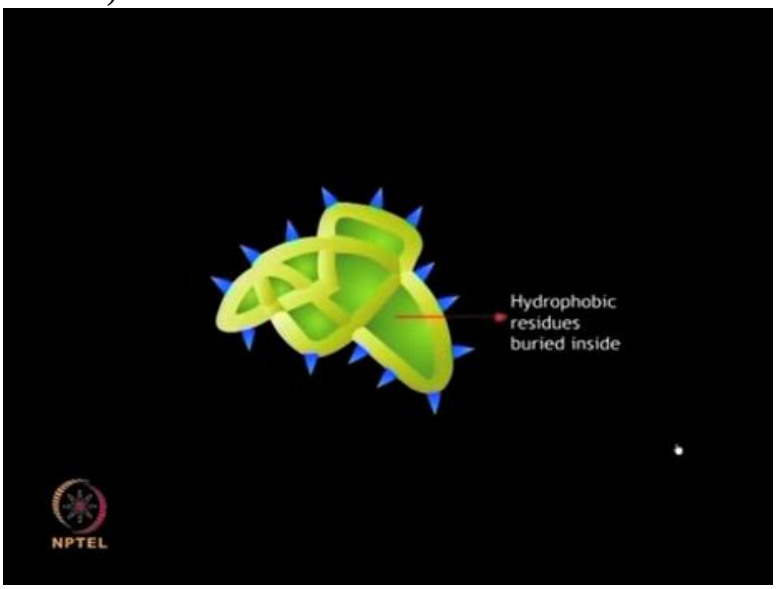
The process of protein folding is dictated by the distribution of polar and non-polar amino acid residues in the protein.

(Refer Slide Time: 12:58)



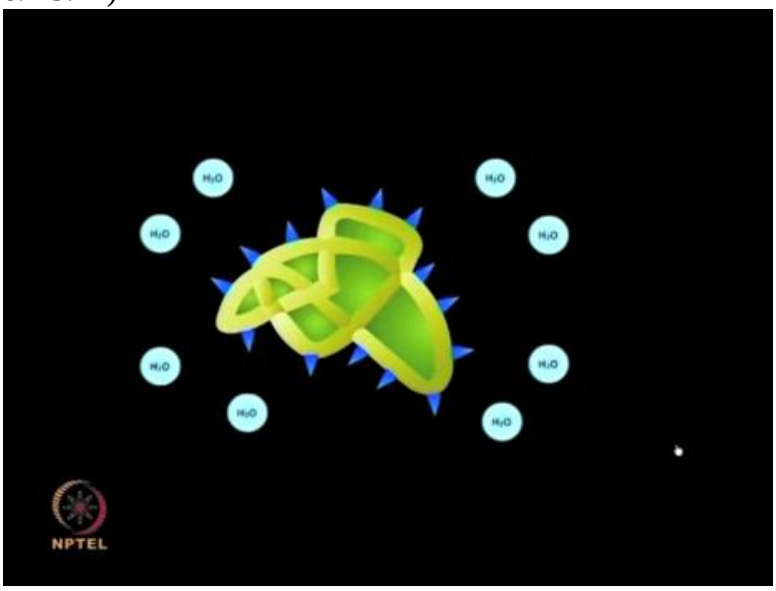
The hydrophobic amino acids are driven to interact with one another…

(Refer Slide Time: 13:05)



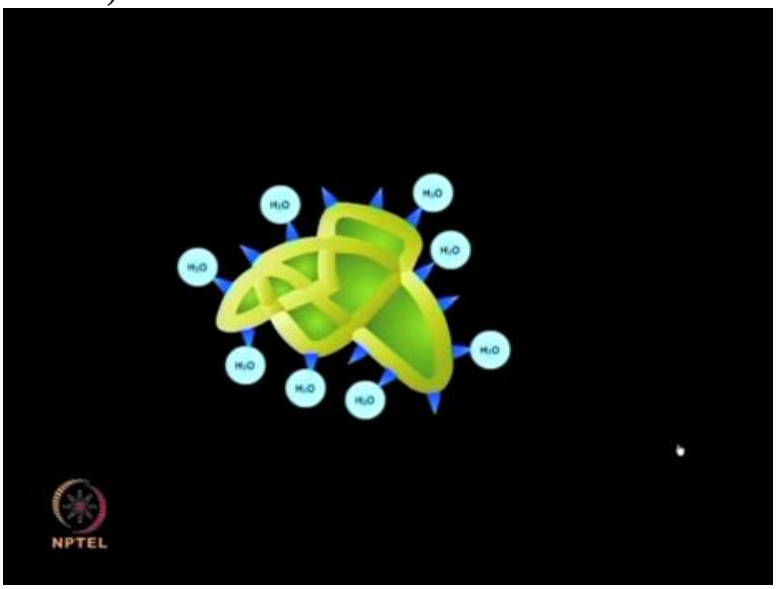
… by a process known as hydrophobic collapse.

(Refer Slide Time: 13:12)



They come together and during this process, eliminate water molecules surrounding them.

(Refer Slide Time: 13:23)

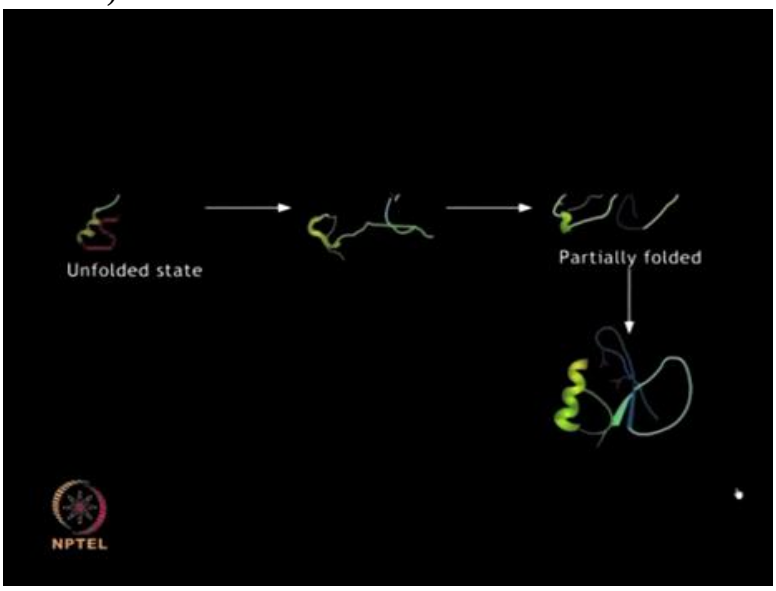


The polar residues remain on the surface and form Hydrogen bonds with water molecules while the hydrophobic residues get buried within the core of protein.

(Refer Slide Time: 13:30)

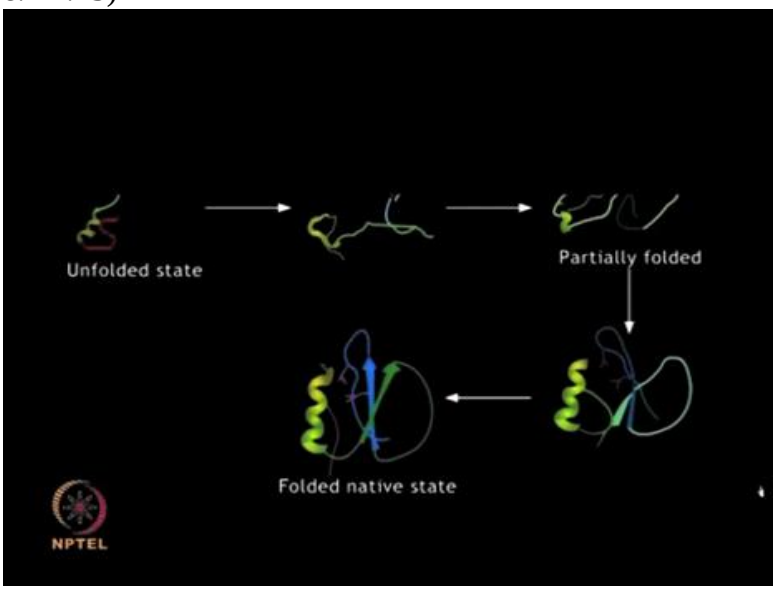
Protein folding is a cooperative process where as the unfolding is a sharp and quick transition. Proteins typically adopt only one characteristic functional native state conformation which has lowest free energy and it is more stable.

(Refer Slide Time: 13:56)



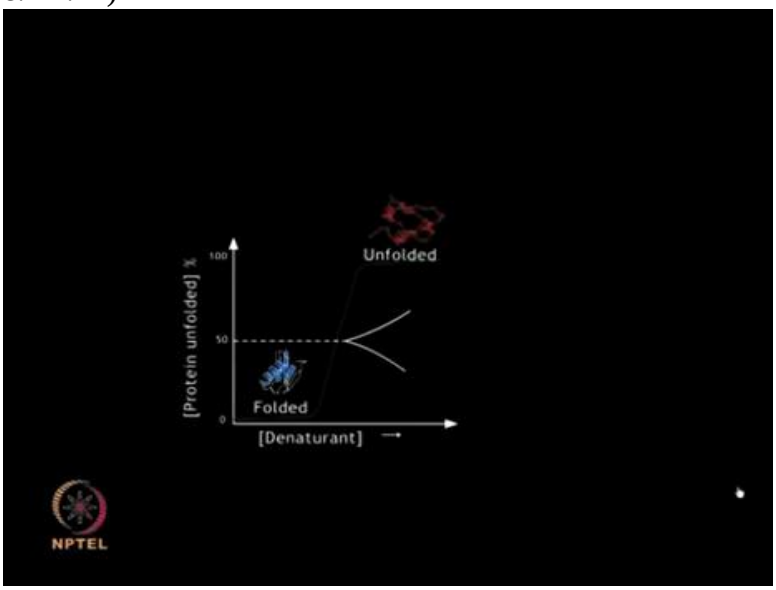
Folding is limited to one conformation due to properties of amino acid side chain such as hydrophobicity, size, shape etc. Folding is highly cooperative process wherein there is progressive stabilization of the intermediates as you can see here.

(Refer Slide Time: 14:25)



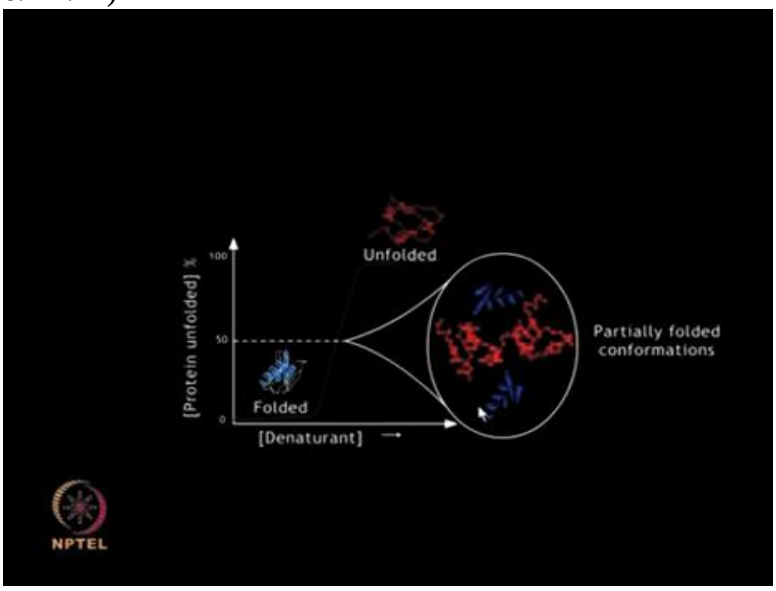
Although it is theoretically possible to predict the protein structure from the amino acid sequence, several long range interactions can often limit such predictions.

(Refer Slide Time: 14:41)



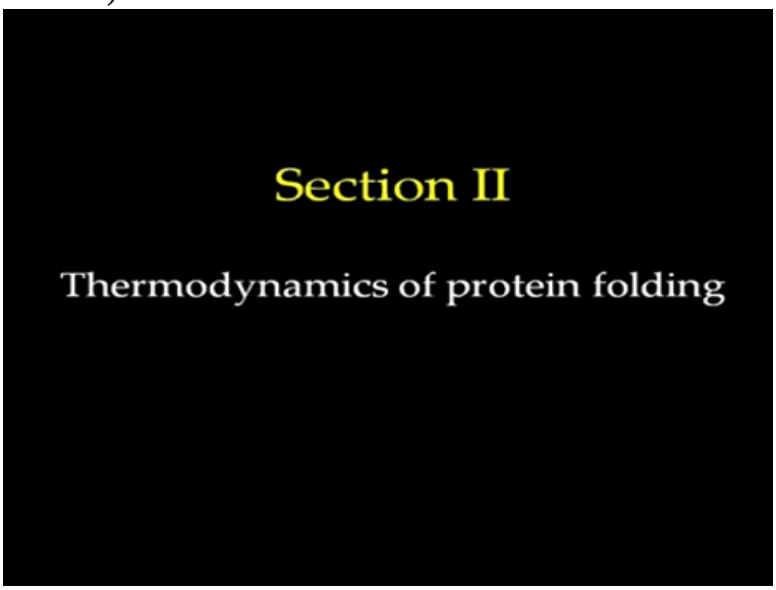
Here on x-axis, denaturants are plotted and y-axis, the percentage protein unfolded is plotted.

(Refer Slide Time: 14:47)

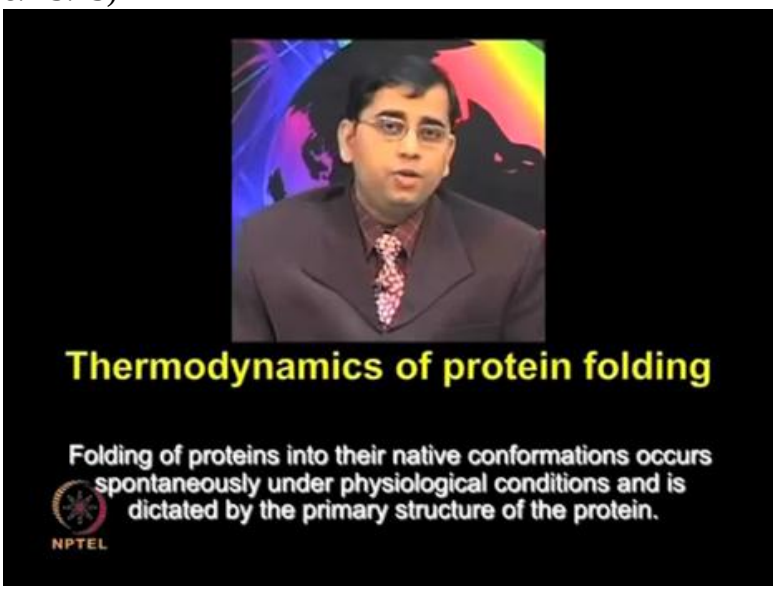


On zero, you can see that is totally folded form of the protein, on 100% it is unfolded form. But if you take a mixture from 50% that is, either unfolded or folded form which shows that protein can assume either folded form or unfolded form of the proteins.

(Refer Slide Time: 15:13)

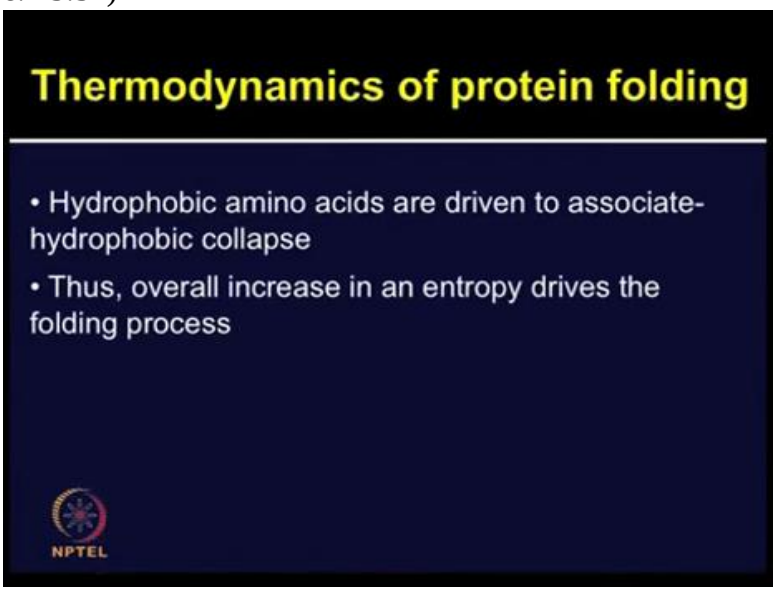


(Refer Slide Time: 15:25)



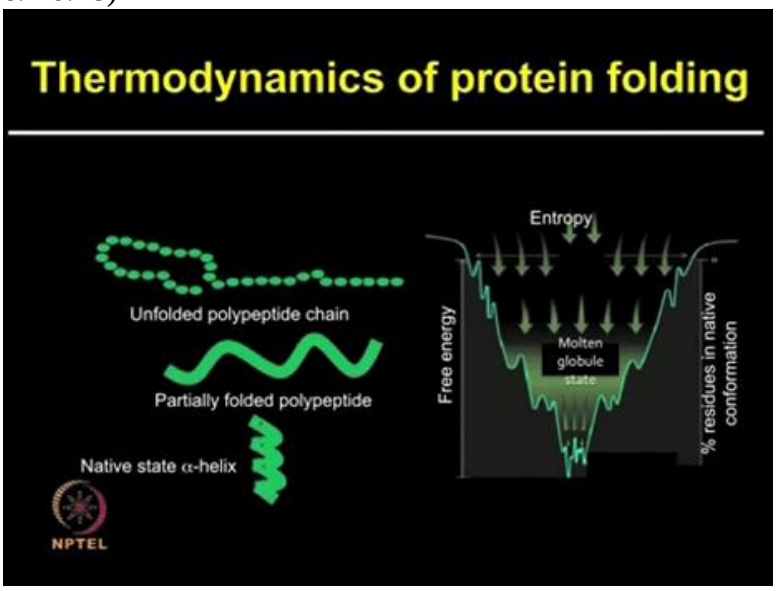
Let's now talk about thermodynamics of protein folding. The folding of proteins into their native conformation occurs spontaneously under physiological conditions and is dictated by the primary structure of protein. Protein folding is thermodynamically favorable process where decrease in free energy from unfolded to folded state occurs.

(Refer Slide Time: 15:54)



Let's talk about some of the basics of thermodynamics for protein folding. As we have seen earlier, the hydrophobic amino acids, they are driven to associate hydrophobic collapse. Therefore overall increase in entropy drives the folding process.

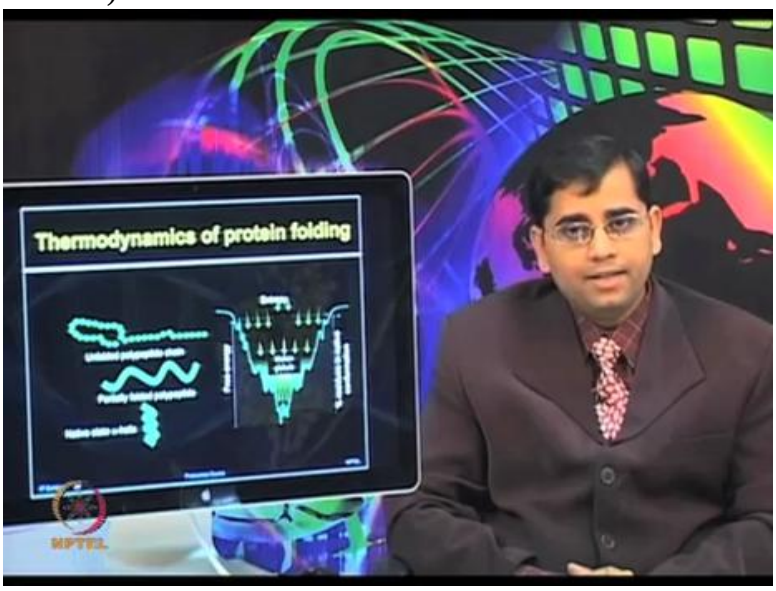
(Refer Slide Time: 16:18)



As you can see in this complex picture here, the folding process can be explained as free energy funnel thermodynamically. If you look at the right-hand side, the open mouth of funnel represents the wide range of structures which are accessible to the ensemble of denatured proteins.

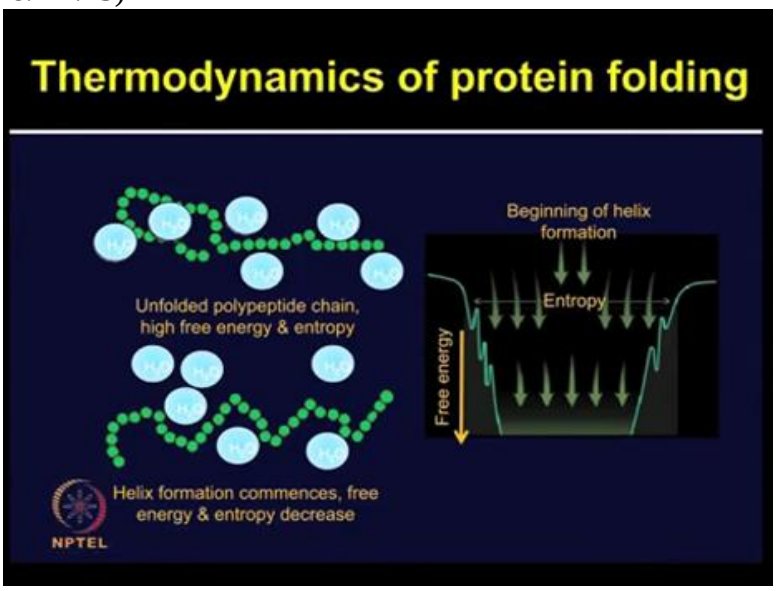
The initial collapsed state of protein with very little thermodynamic stability is known as molten globule.

(Refer Slide Time: 17:03)



The amino acid side chains are extremely disordered in this state and several fluctuations can be observed as you can see from these arrows. As free energy of protein molecules decreases, the protein molecules move down to the narrower part of the funnel.

(Refer Slide Time: 17:45)

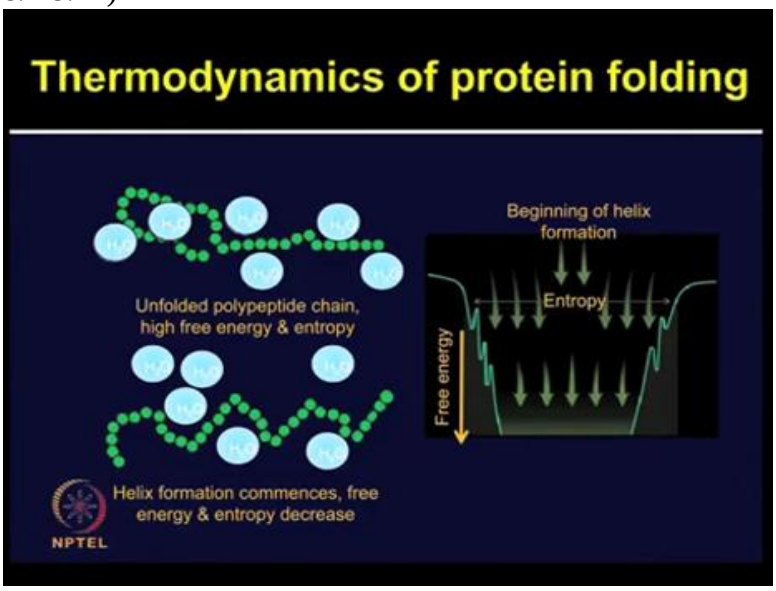


Look at the bottom part here and only few conformations can be accessible here. So at the bottom of the funnel, well-defined and folded conformation states are present. So if you look at unfolded polypeptide chain, so the amino acids that have been joined together by the peptide bonds but they have not yet formed their secondary or tertiary structure.

So this conformation has highest free energy and entropy. The amino acids in the polypeptide chain start interacting by means of hydrogen bonds across the polypeptide backbone in order to initiate the folding process. The free energy and entropy of the system gradually decreases as folding takes place.

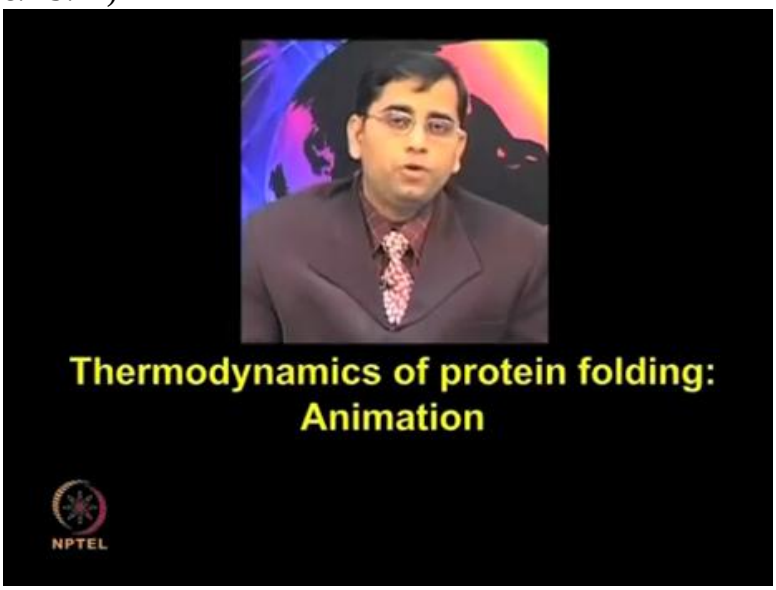
The entropy of the polypeptide chain decreases during this process. So in the thermodynamic terms, lowering of entropy is favored by a corresponding increase in entropy in the surroundings composed of the water molecules.

(Refer Slide Time: 18:42)



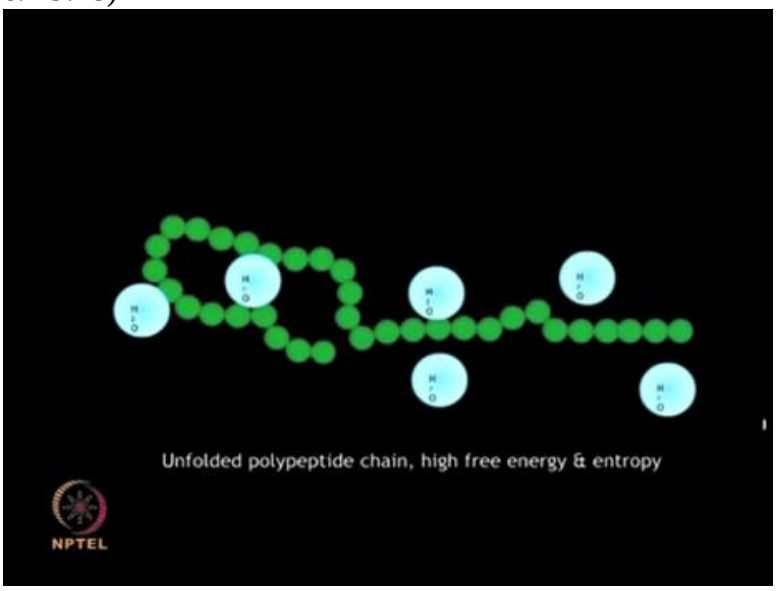
Now if you look at the funnel again, as the protein continues to fold in order to assume its stable low energy native state conformations, then entropy also decreases. While it may appear unfavorable for the system, however entropy of the surrounding water molecules increases in this process and it increases overall entropy and makes it favorable and spontaneous.

(Refer Slide Time: 19:17)



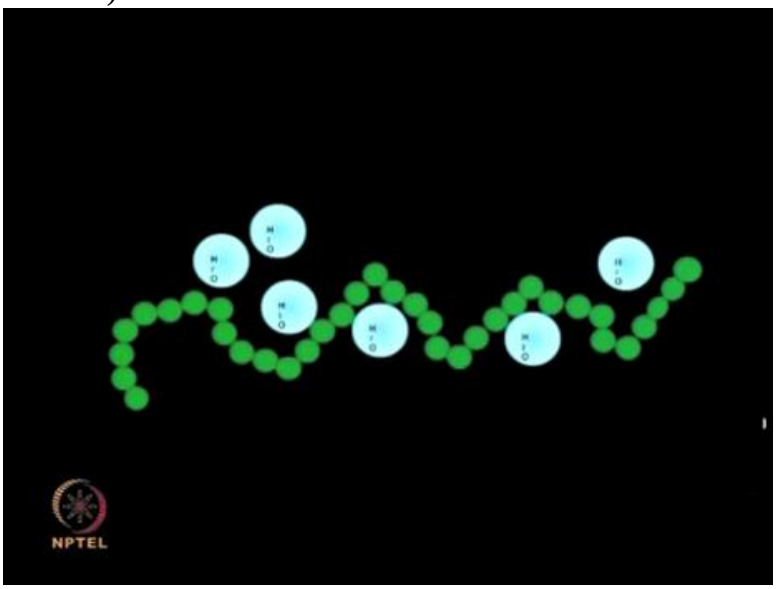
So let me show you how protein folding works and how it can be described in the thermodynamic terms in following animation.

(Refer Slide Time: 19:28)



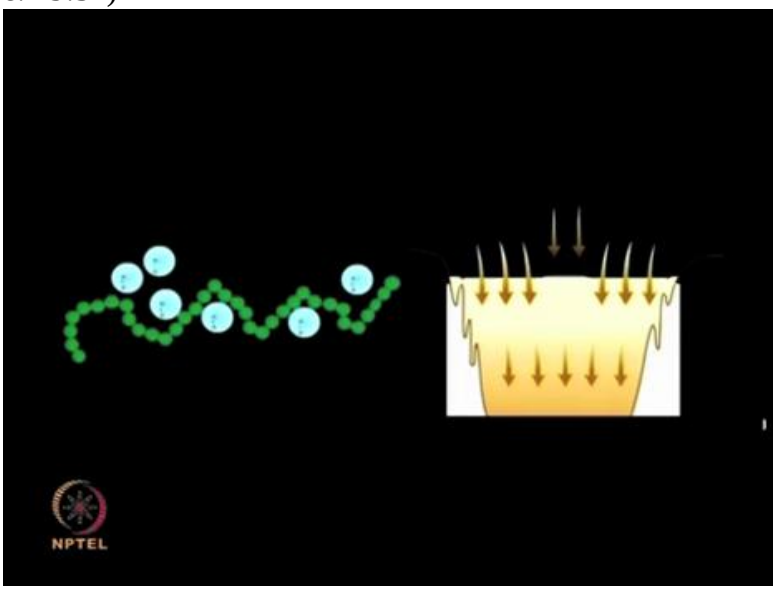
An unfolded polypeptide chain has very high energy and entropy.

(Refer Slide Time: 19:45)



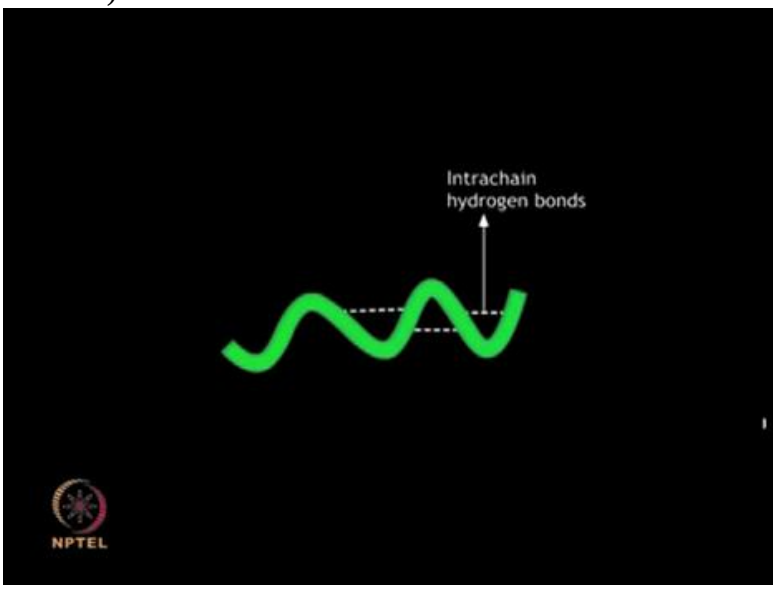
The protein folding acts to decrease the free energy of the system by forming favorable interactions and assuming a more stable state.

(Refer Slide Time: 19:51)



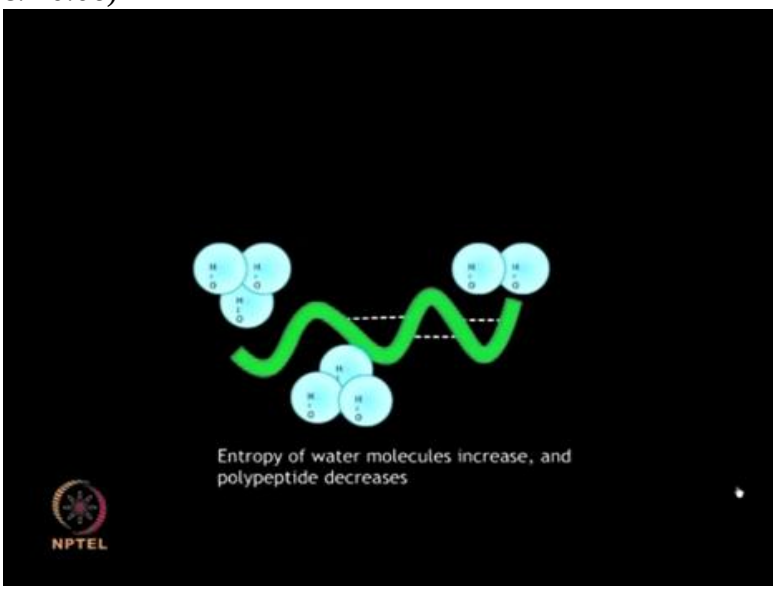
The entropy of the polypeptide chain decreases during this process.

(Refer Slide Time: 19:54)



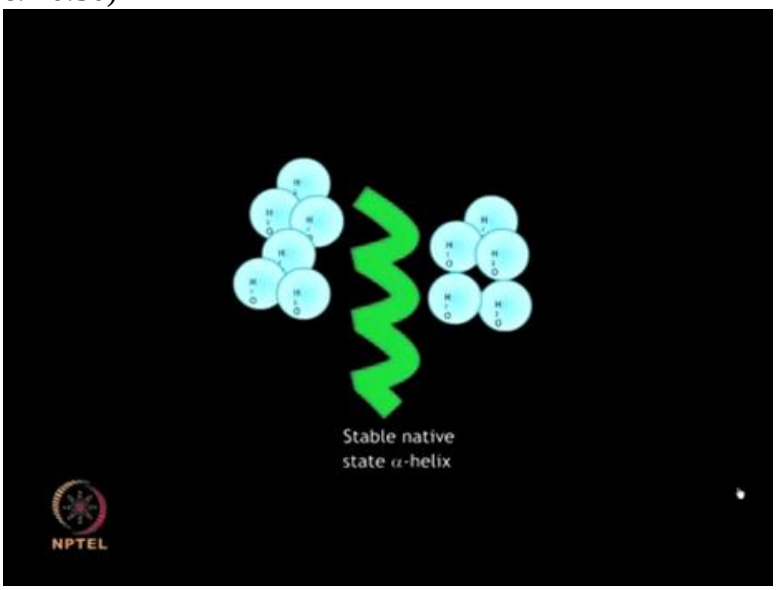
As protein contains to fold in order to assume its stable low energy native state conformation, the entropy also decreases.

(Refer Slide Time: 20:06)



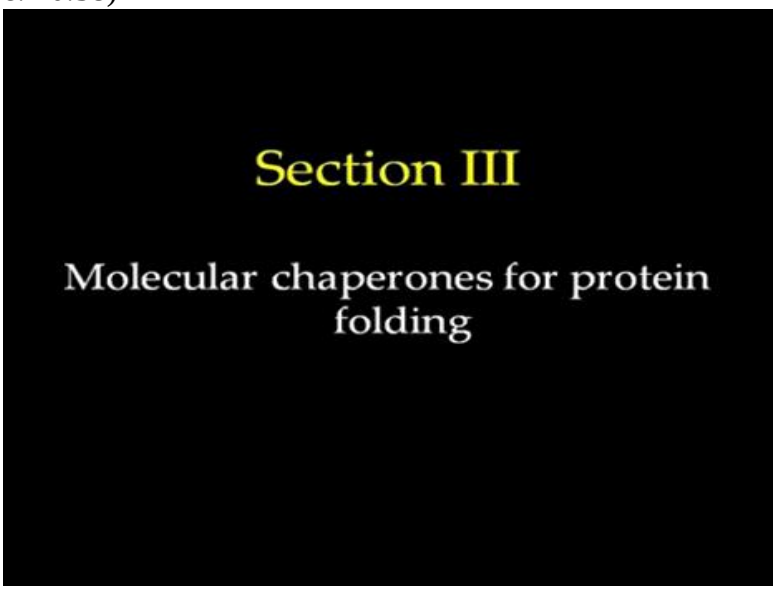
As the protein continues to fold in order to assume its stable, low energy native state conformation, the entropy also decreases.

(Refer Slide Time: 20:30)



While this should seem unfavorable for the system, it must be recalled that the entropy of the surrounding water molecule increases during the process thereby increasing overall entropy and making it favorable and spontaneous.

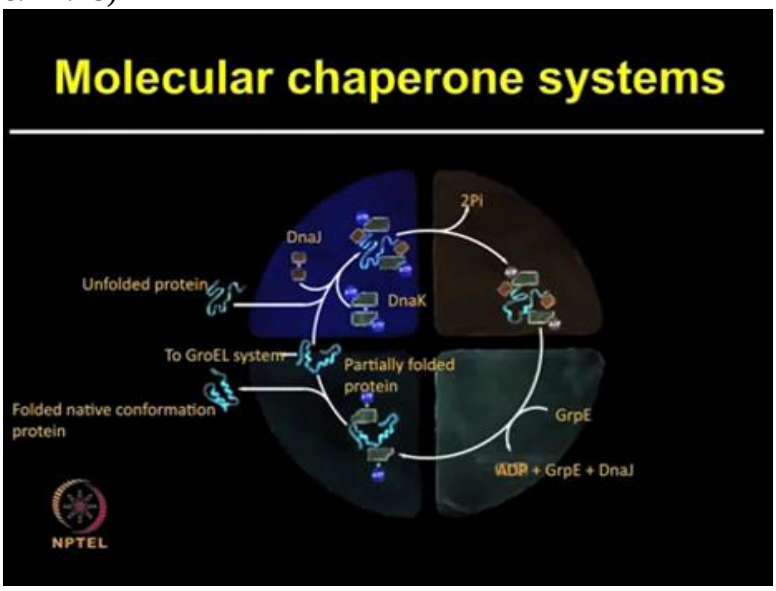
(Refer Slide Time: 20:38)



(Refer Slide Time: 20:44)

Let's now talk about molecular chaperones for protein folding. The molecular chaperones are class of heat inducible proteins which provides kinetic assistance in protein folding process. They prevent protein aggregation and promote protein folding by binding to the hydrophobic surfaces which are exposed in non-native protein conformations.

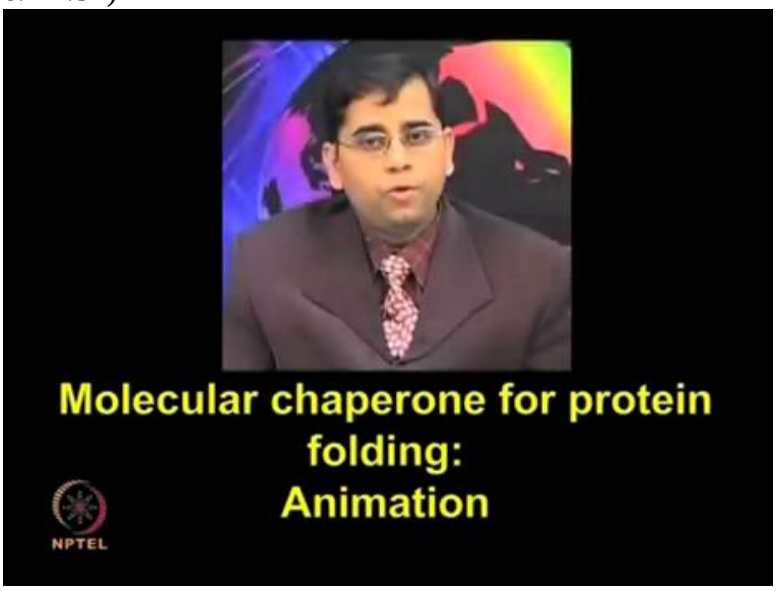
(Refer Slide Time: 21:18)



So let's talk about various molecular chaperone systems. Many newly synthesized proteins form folded structures in-vivo spontaneously and without any assistance. However folding efficiency could be limited by various processes such as protein aggregation which are promoted by the transiently exposed hydrophobic surfaces.

In response to the heat shock, the cells produce significant amount of unfolded proteins by synthesizing new systems which are known as molecular chaperones which are designed to promote the protein folding process. There are several molecular chaperone systems which have been exemplified in E. coli, include GroEL GroE system, DnaK, DnaJ, GrpE and ClpB.

(Refer Slide Time: 22:34)

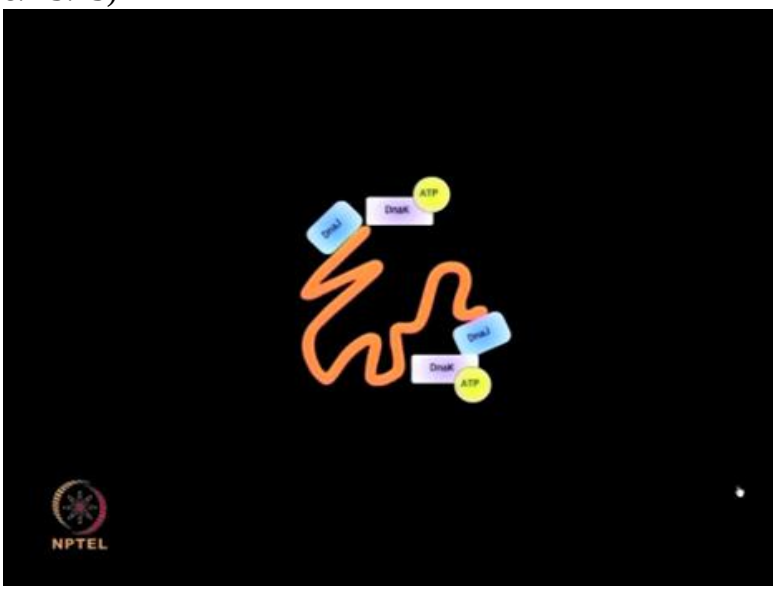


The bacterial chaperone GroEL, it binds proteins in non-native state and allows enzymes to be recovered quantitatively in native form by binding which requires co-chaperinin, GroES, and ADP. Let me show how this chaperinin works and governs the protein folding process in following animation.

(Refer Slide Time: 23:06)

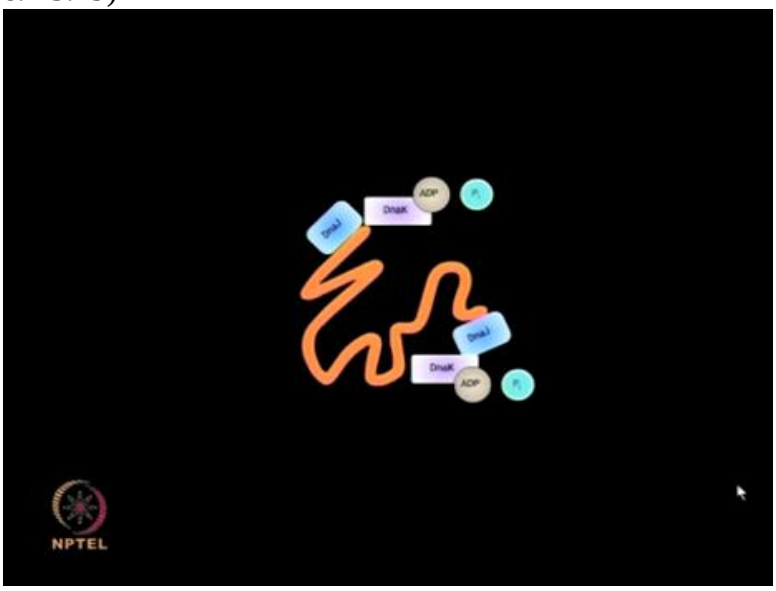
The unfolded protein is bound by DnaJ and…

(Refer Slide Time: 23:13)



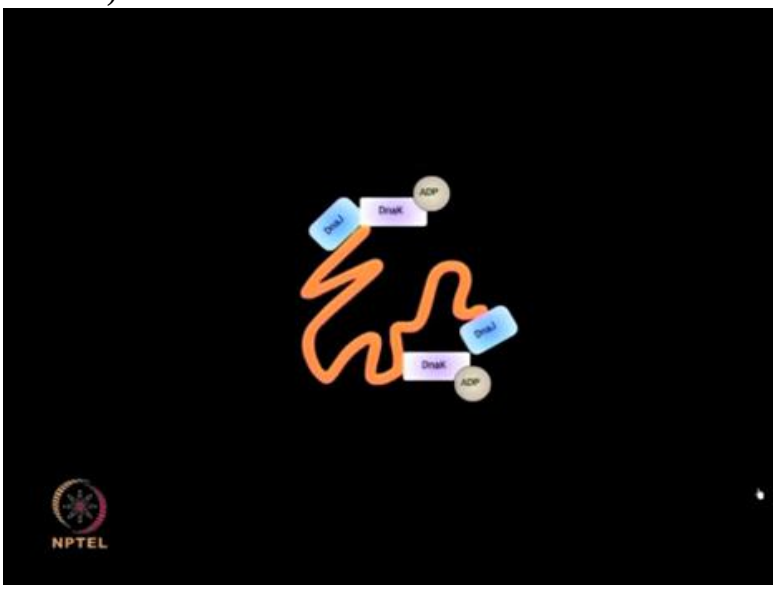
…then by DnaK which is an ATP bound protein.

(Refer Slide Time: 23:19)



The hydrolysis of ATP into ADP and Pi by DnaK is stimulated by DnaJ.

(Refer Slide Time: 23:33)

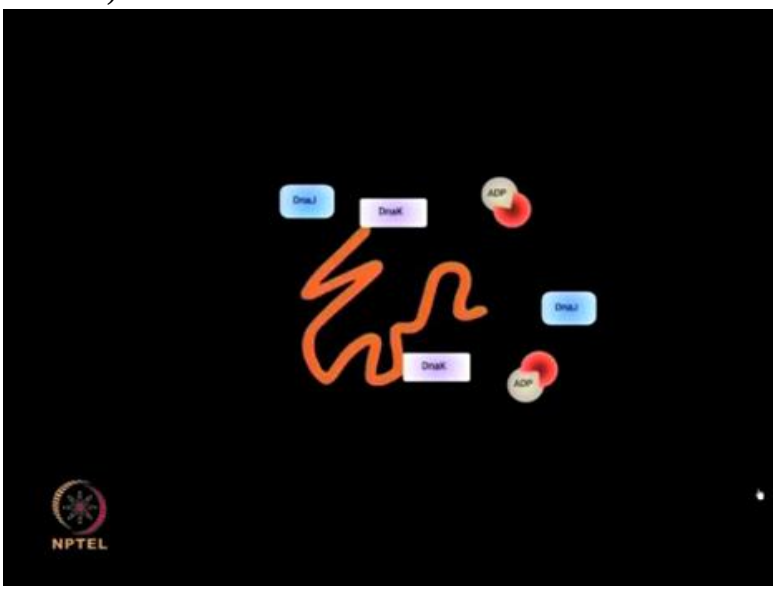


This resulting DnaK ADP remains tightly bound to the unfolded protein.

(Refer Slide Time: 23:38)

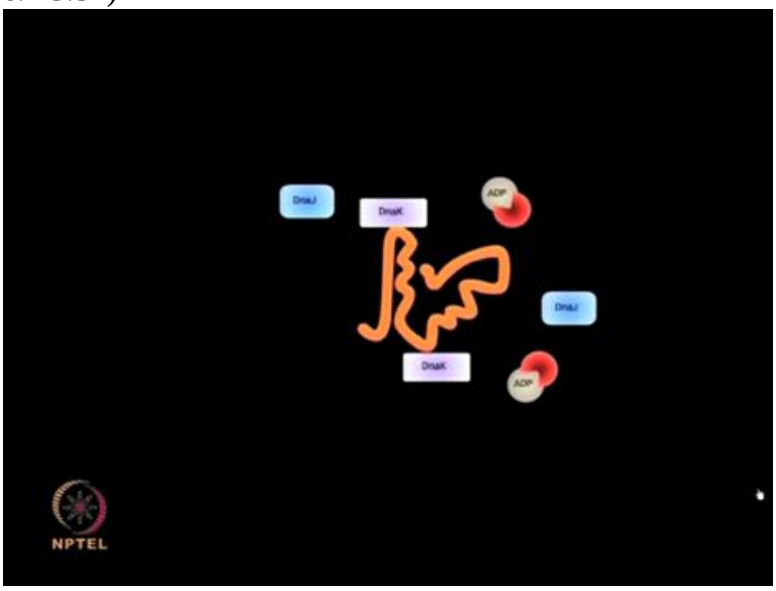
The nucleotide exchange factor GRPE present in bacteria…

(Refer Slide Time: 23:49)



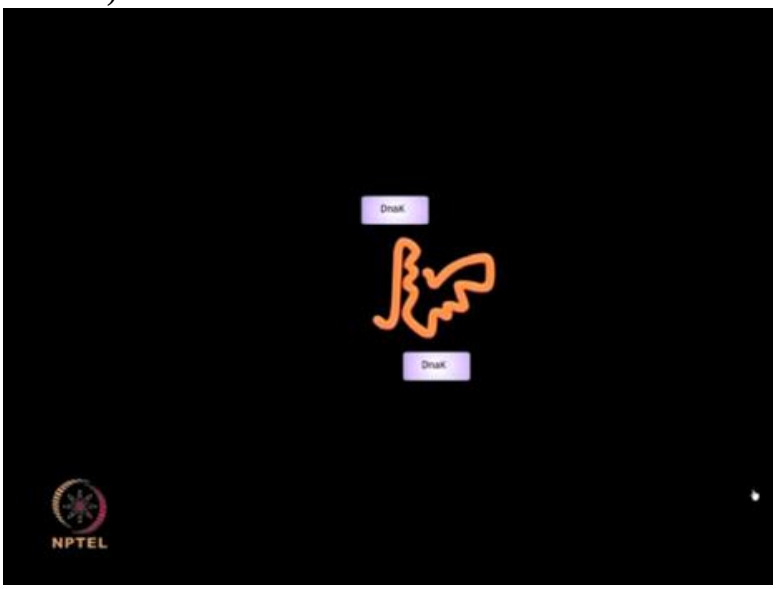
… facilitates release of ADP along with DnaJ.

(Refer Slide Time: 23:52)



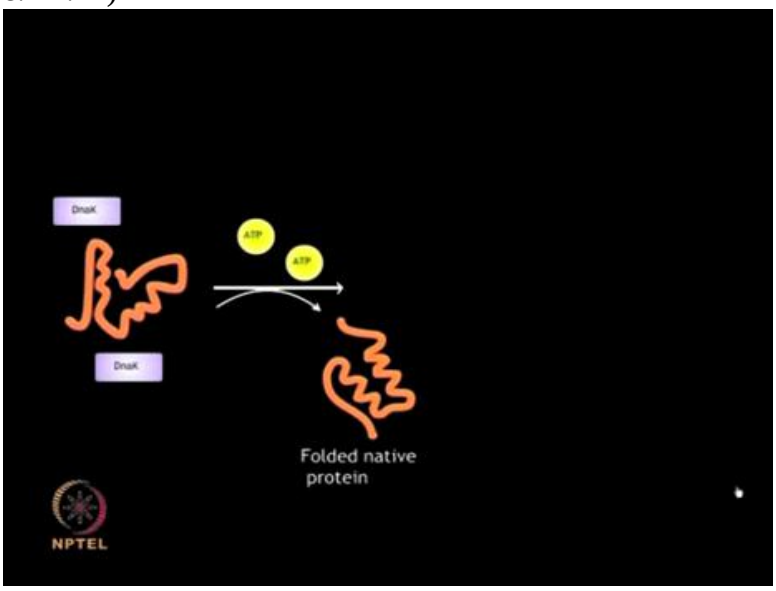
This leaves the DnaK bound to the partially folded protein …

(Refer Slide Time: 23:59)



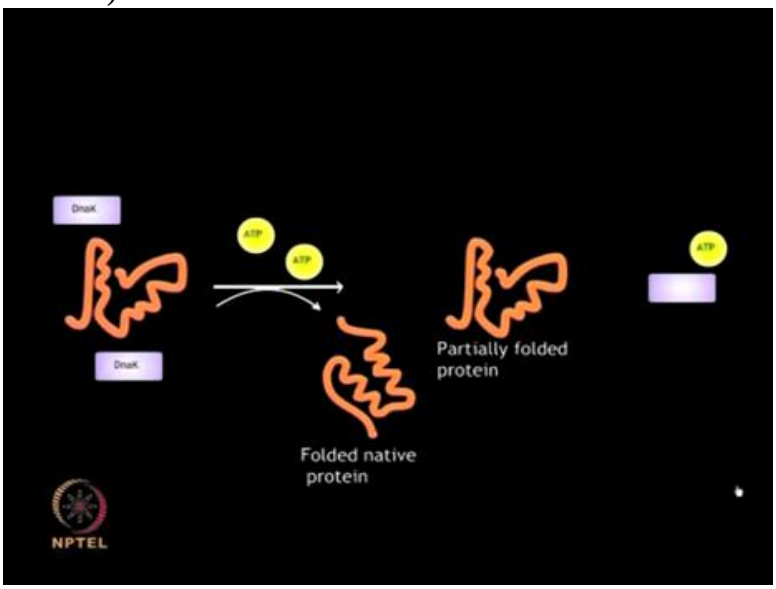
… which continues to undergo folding to a more favorable low energy conformation. Once the protein gets completely folded, it gets detached from DnaK…

(Refer Slide Time: 24:14)



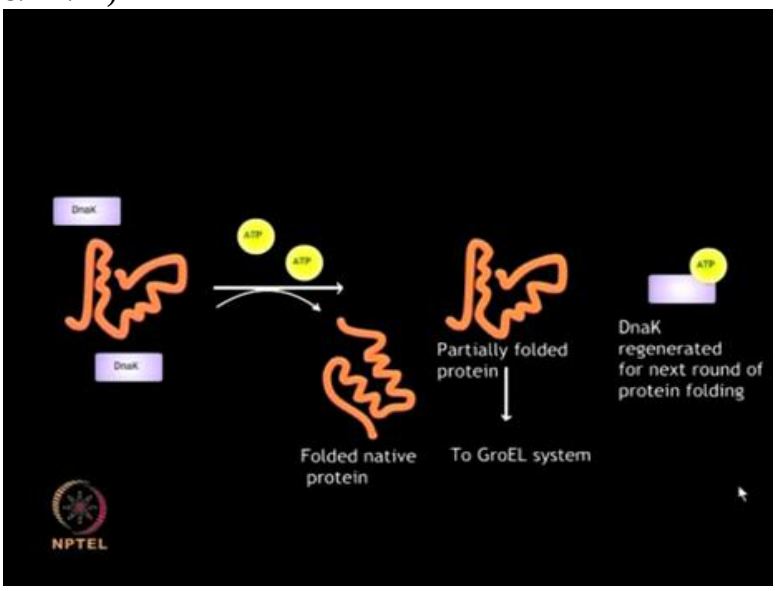
…which then binds ATP again…

(Refer Slide Time: 24:19)



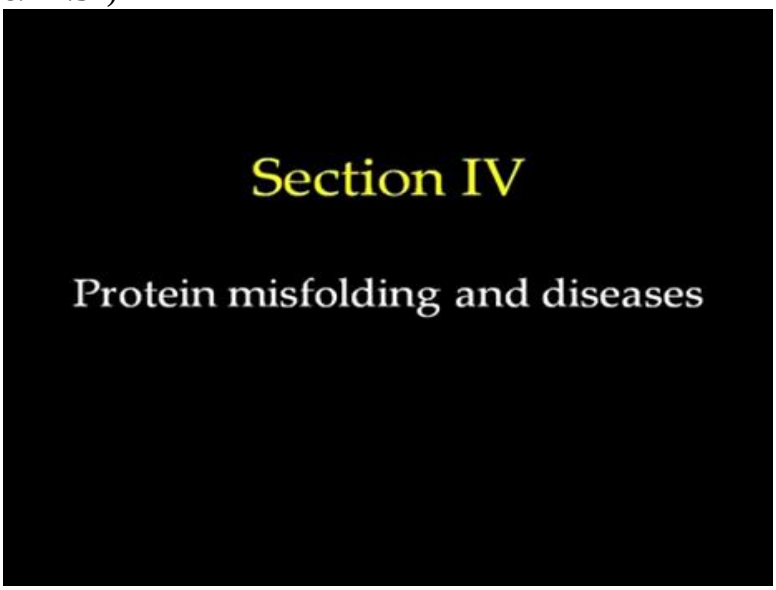
…and completes the cycle and prepares it for next round of protein folding.

(Refer Slide Time: 24:24)

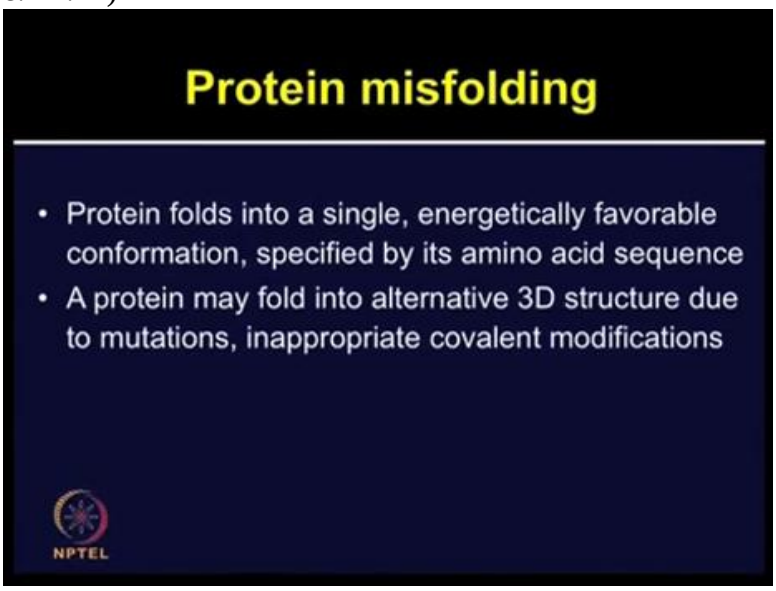


Any protein which may not have been folded completely is then taken over by the GroEL chaperone system which completes the fold.

(Refer Slide Time: 24:34)



(Refer Slide Time: 24:41)

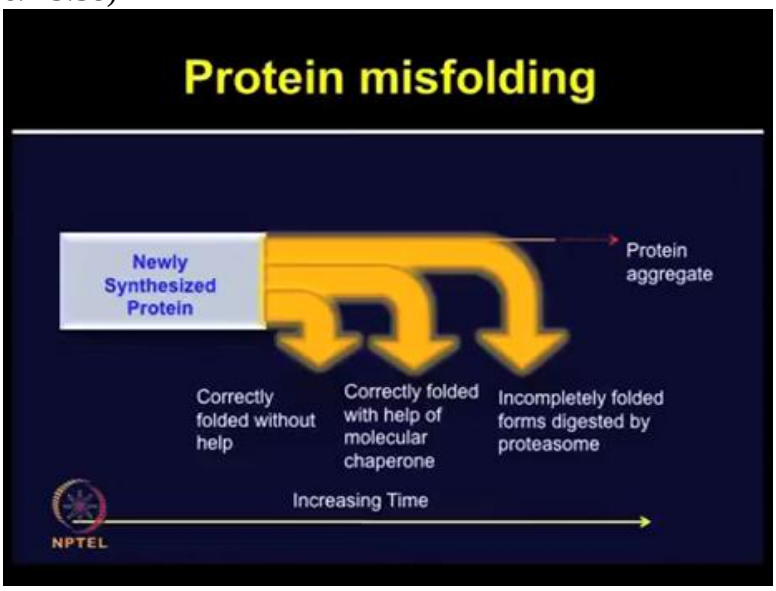


After talking about how protein folding works, lets discuss about protein misfolding and how misfolding may result into various diseases. So protein misfolding results into large number of human diseases which arise as a consequence of protein misfolding.

In protein folding, mutations cause defective folding, aberrant assembly and incomplete processing which results into altered folding properties. Proteins fold into a single energetically most favorable conformation which is specified by its amino acid sequence.

A protein may fold into alternative three-dimensional structures because of mutations or inappropriate covalent modifications. Therefore protein misfolding may lead to loss of normal protein function.

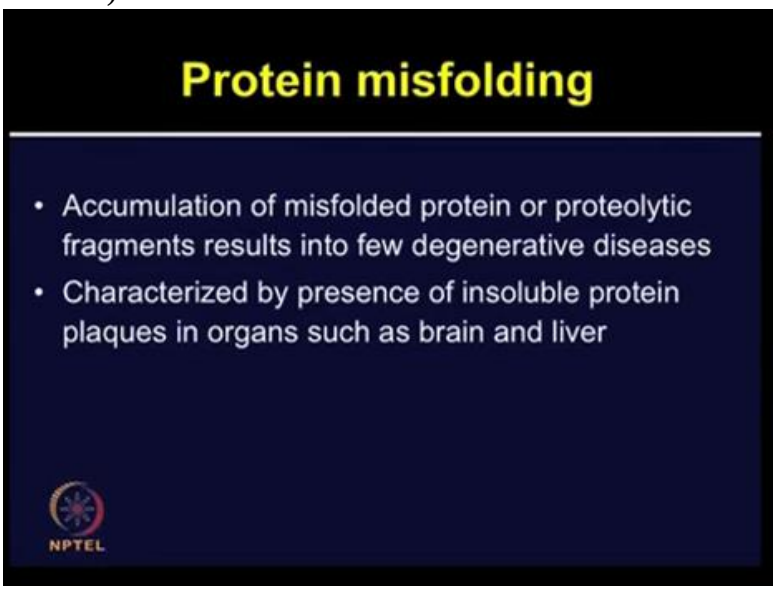
(Refer Slide Time: 25:58)



In this slide, we can see from a newly synthesized protein; it could have various fate. It may form the proper folded form without any assistance or it can give correct folded forms in presence of molecular chaperones or incompletely folded forms can be digested by proteasome machinery or it may result into the protein aggregation.

So a newly synthesized protein may give rise to any of these forms depending on various factors which are going to govern the protein folding process.

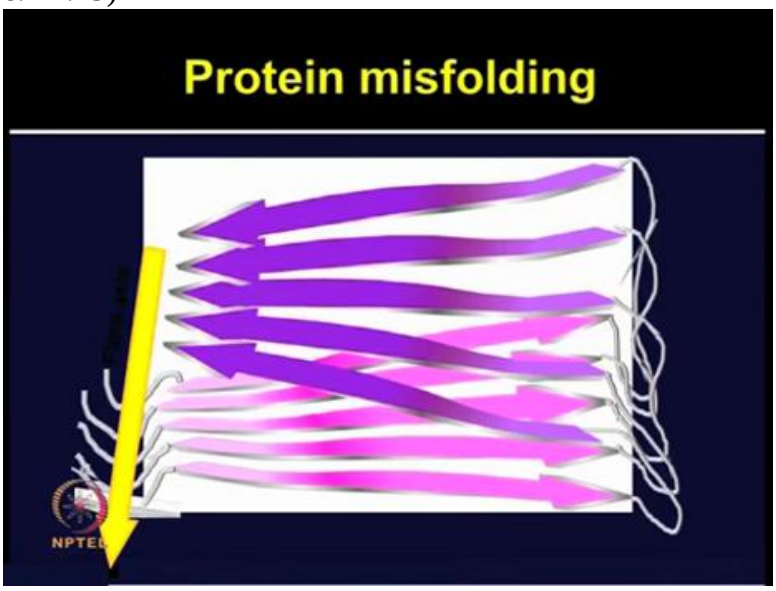
(Refer Slide Time: 26:39)



So accumulation of these misfolded proteins or proteolytic fragments can result into few degenerative diseases. These degenerative diseases are characterized by the presence of insoluble protein plaques in organs such as brain and liver.

For example, Alzheimer's disease in human as well as Parkinson's disease, bovine spongiform encephalopathy also known as mad cow disease in cows and Scrapie disease in sheep.

(Refer Slide Time: 27:19)

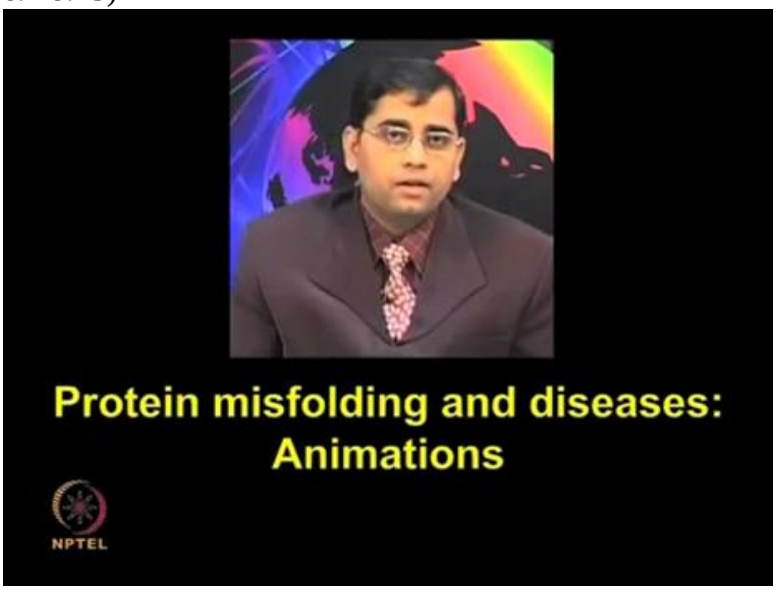


In this slide, it is shown that amyloid fibers are involved in neuro-degenerative diseases and the protein aggregation due to the large beta sheets are deduced from solid state NMR.

In Alzheimer's disease, the presence of beta amyloid containing plaques is associated with neuro-degeneration and dementia. In other neuro-degenerative diseases, it has also been shown that it involves protein aggregation.

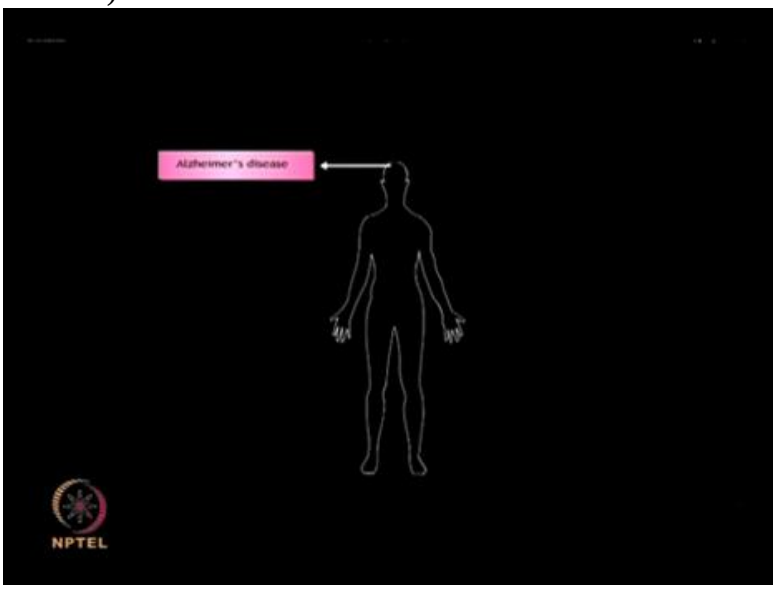
Prion diseases such as Jacob's disease and BSE or bovine spongiform encephalopathy are associated with amyloid deposit of PrP proteins. So how insoluble protein aggregates can result into various diseases?

(Refer Slide Time: 28:23)



Let’s discuss some of protein misfolding related diseases in following animation.

(Refer Slide Time: 28:41)

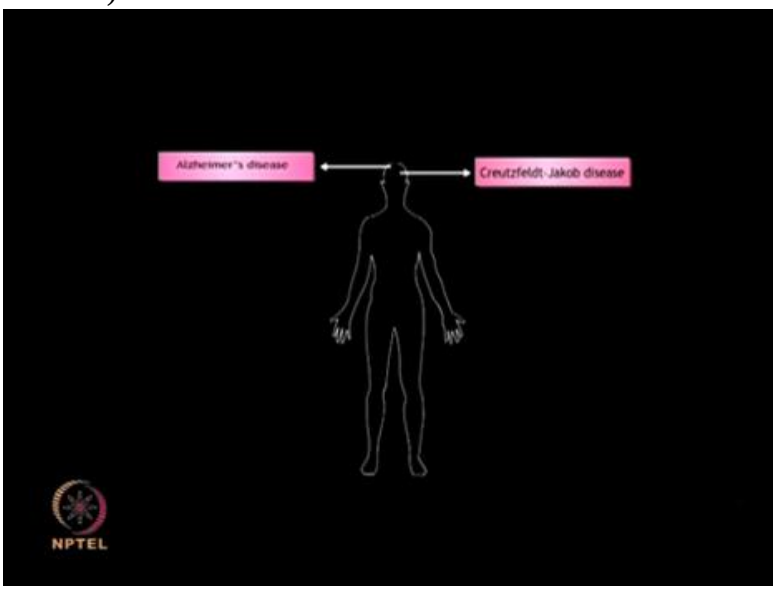


Protein misfolding results into various diseases such as Alzheimer’s disease. In Alzheimer's disease, the structure of certain normal soluble, cellular proteins which are normally rich in alpha helical regions are converted into beta strand conformations which further link with each other to form beta sheet aggregates known as amyloids.

The insoluble amyloid plaques are essentially made up of a single polypeptide chain or fibrals known as amyloid B protein. It is observed in the brain of the patients with Alzheimer's where dead or drying neurons are surrounded with plaques.

The neuro-toxicity is believed to be caused by the amyloid fibrals before they get deposited as amyloid plaques. This disease presents various symptoms such as memory loss, decreased neuro-muscular coordination, confusion and dementia.

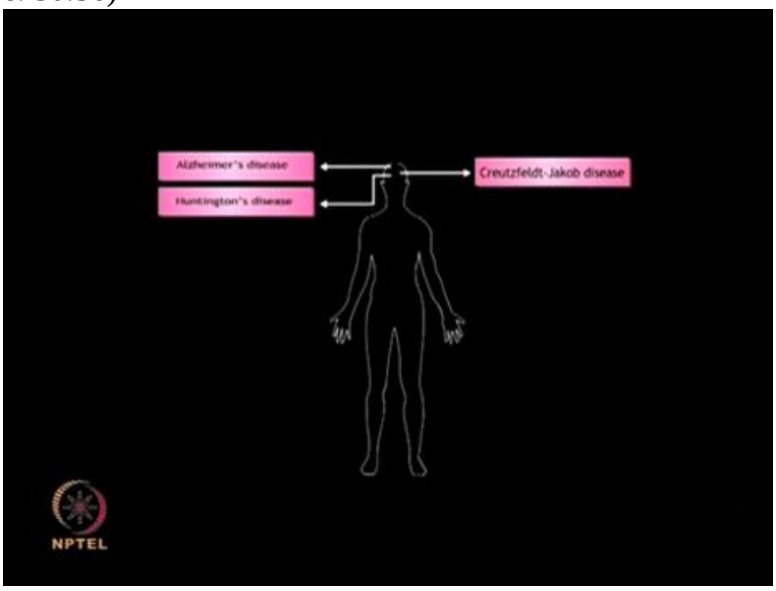
(Refer Slide Time: 30:07)



Jacob's disease: It was initially believed to be caused by viruses or bacteria. However later, it was discovered to be transmitted by small proteins known as prions.

The prion proteins are composed of beta sheet structures that have been modified from previously existing alpha helices. The protein aggregates of one abnormal protein is sufficient to function as nuclei for other normal proteins to attach.

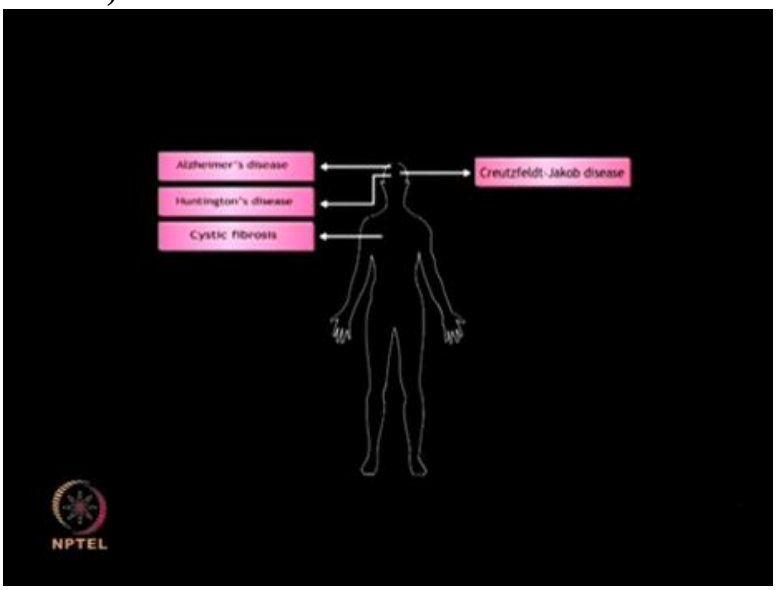
(Refer Slide Time: 30:56)



Huntington's disease: It is a neuro-degenerative disorder of genetic origin which affects muscular coordination. It is caused by increased number of trinucleotide repeats CAG in Huntingtin gene leading to increased number of glutamine residues incorporated in corresponding protein.

This alters the folding of Huntingtin protein which has highest concentration in brain and testes. The exact function of this protein is unclear but it is known to interact with several other proteins. The mutative protein has also been found to have effects on chaperone proteins which in turn help to fold several other proteins.

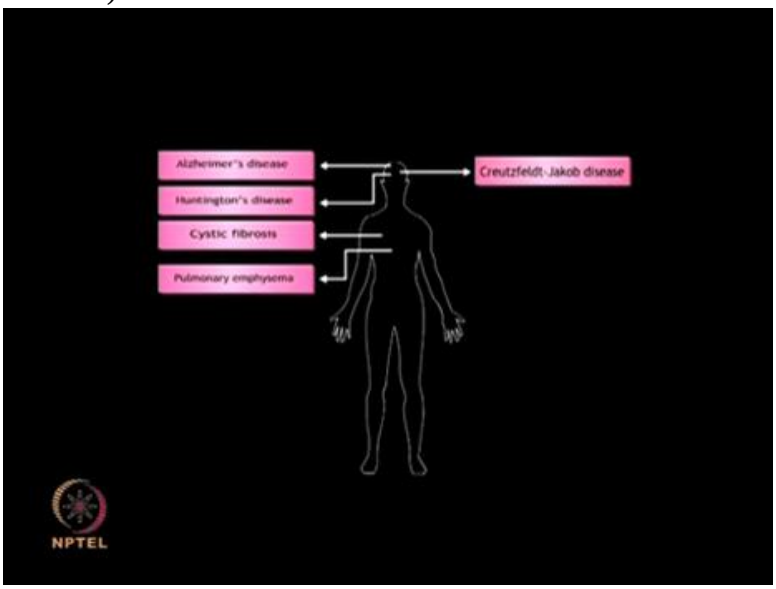
(Refer Slide Time: 32:01)



Cystic Fibrosis: This is an autosomal recessive disorder caused by mutation in gene for the protein Cystic Fibrosis Transmembrane conductance Regulator or CFTR. The CFTR gene regulates components of sweat, digestive juices, and mucus.

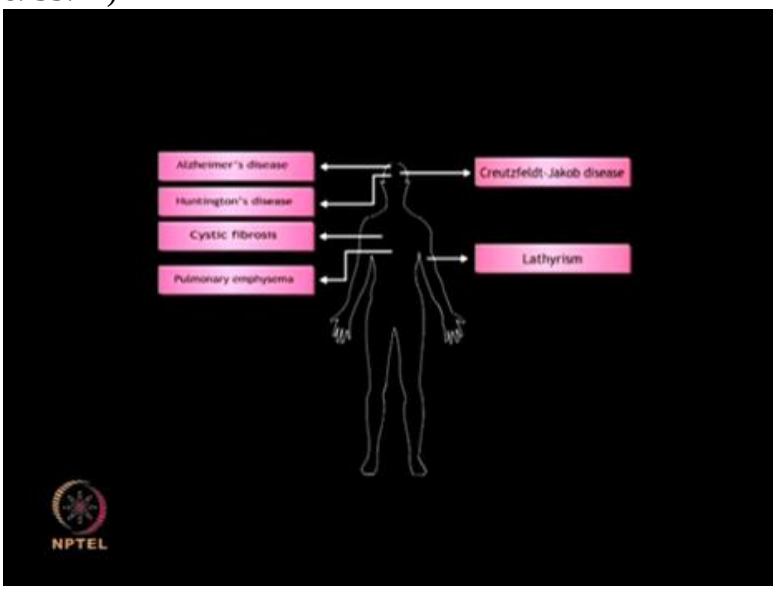
It is caused by deletion of 3 nucleotides leading to the elimination of phenylalinine residue from the protein and therefore results into abnormal folding. The dysfunctional protein gets degraded by the cell.

(Refer Slide Time: 32:49)



Pulmonary emphysema: It is a progressive disease of lungs which causes shortness of breath. It can be caused by deficiency of protein Alpha-1 AntiTrypsin or A1AT. The A1AT gene is responsible for protection of lung tissue from damage by enzyme neutrophil elastase. Abnormally secreted A1AT gets accumulated in liver thereby allows lung tissue damage. The disease results into wheezing, shortness of breath and asthma like symptoms.

(Refer Slide Time: 33:44)

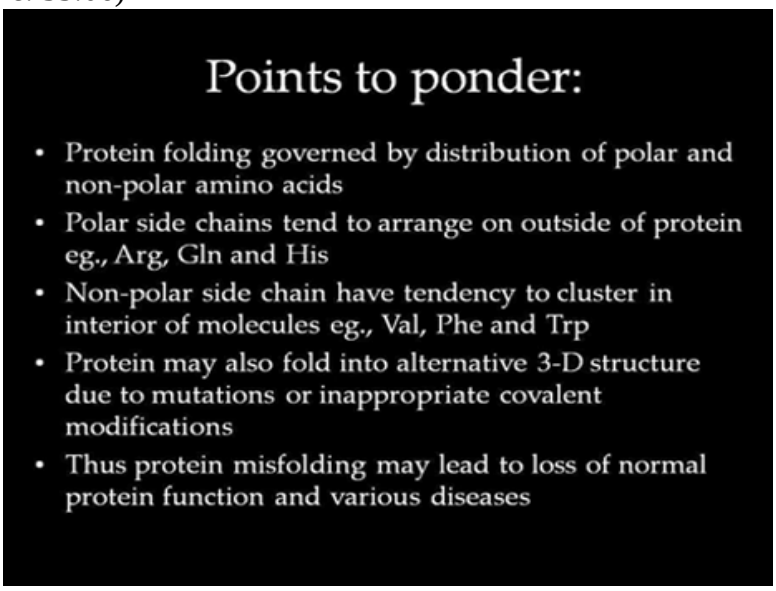


Lathyrism: It is regular ingestion of seeds from sweet pea Lathyrus odoratus which causes disruption of cross-linking in the muscle protein Collagen. Collagen is very important structure protein which has triple helical structure.

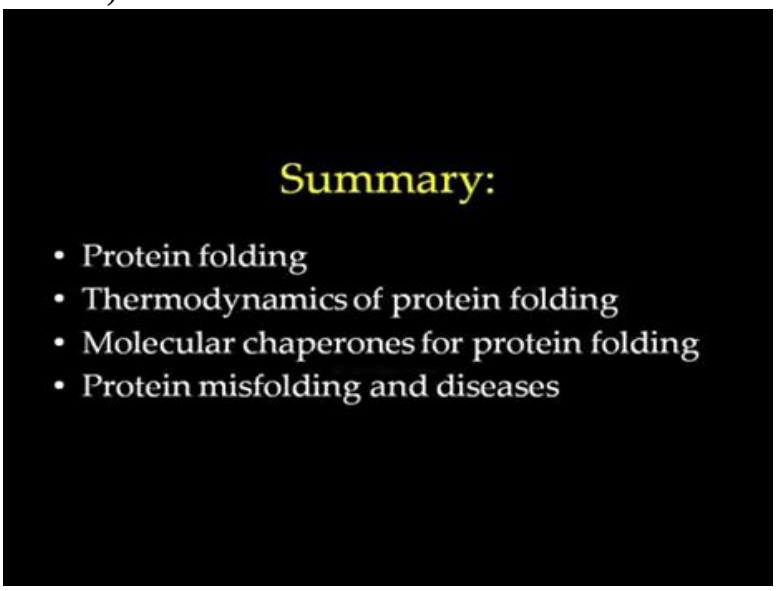
The cross-links formed are due to the oxidation of Lysine residues by the enzyme Lysyl oxidase to form Allysine. These are essential for proper folding of Collagen and giving it the required strength. B aminoproprionitrile present in abundance in sweet pea deactivates this enzyme by binding to its active site.

This prevents cross-linking and proper folding of the protein. It may also result in muscle fragility and weakness.

(Refer Slide Time: 35:06)



(Refer Slide Time: 35:12)



In summary, we talked about protein folding and how various polar and non-polar side chains restrict and govern the process of protein folding. We then looked at the thermodynamics of protein folding albeit very briefly. We talked about entropy and how it governs the protein folding.

The molecular chaperones, we talked about some classical examples in animations and then we discussed about protein misfolding and described some of the diseases which may result due to the protein misfolding. We will continue our discussion about basics of protein structure and function in next class as well. Thank you