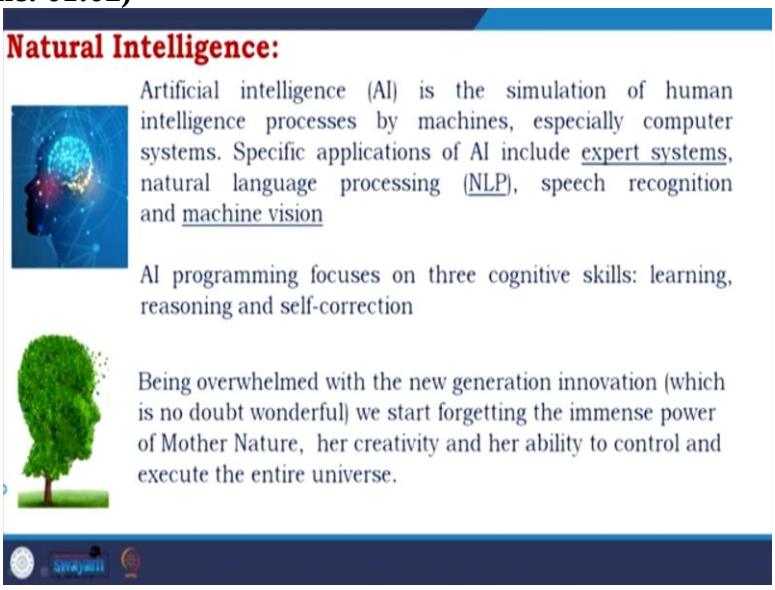
Structural Biology
Prof. Saugata Hazra
Department of Biotechnology
Indian Institute of Technology, Roorkee

Lecture - 09
Protein - Super Secondary Structures, Motif, Domains, Non-Covalent Interactions

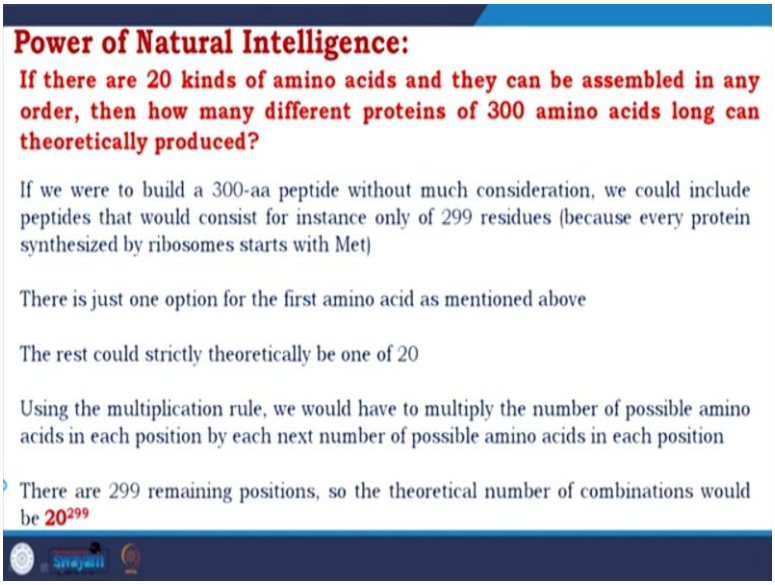
Hi everyone, again, welcome to the course on structural biology. We are continuing with the protein module. And today, I will teach you about super secondary structures, motif, domains, their differences, non-covalent interactions, which play a critical role in protein folding. Before going to super secondary structures, I want to remind you about something that people forget when we are in artificial intelligence. What is artificial intelligence?

(Refer Slide Time: 01:02)



It simulates human intelligence processes by machines, especially computer systems. Specific applications of AI include expert systems, natural language processing (NLP), speech recognition, and machine learning. AI programming focuses on three cognitive skills learning, reasoning, and self-correction. Whereas in the current era, we have seen a lot of progress in machine learning in artificial intelligence. That needs technical advancement that needs appreciation, and they are doing good. Being overwhelmed with a new generation innovation that is undoubtedly wonderful and is no doubt in its appreciation, we start forgetting the immense power of Mother Nature, her creativity, and her ability to control and execute the entire universe.

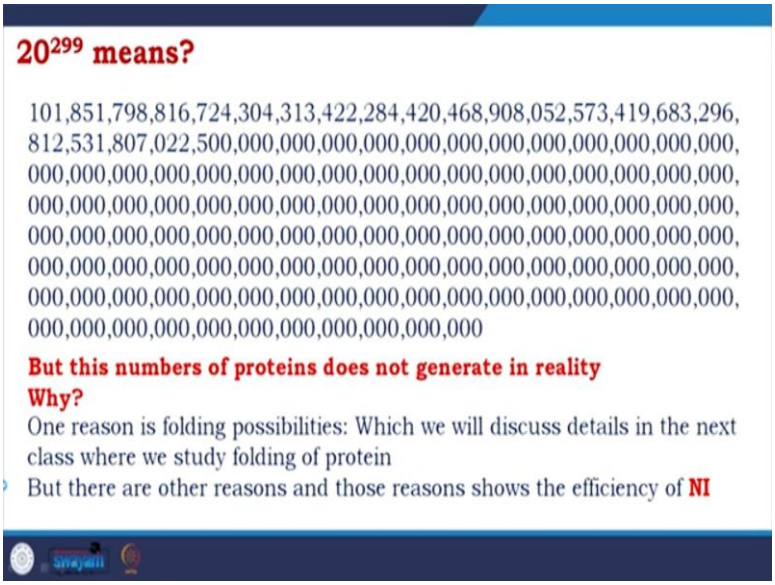
(Refer Slide Time: 02:48)



Suppose there are 20 amino acids, as we know, already from our course. In that case, there are 20 amino acids, and they can be assembled in any order. Then how many different proteins have 300 amino acids long can theoretically be produced. What do we have to do to build a 300 acid peptide? The first position would always be methionine. So in that position is fixed that it would be methionine. Other 299 positions, there would be anything. So if we were to build a 300 amino acid peptide, without much consideration, we would include peptides that would consist, for instance, only of 299 residues because every protein synthesized by ribosome starts with methionine. So you have 20 amino acids, and you have to fill up 299. So the answer is the number of proteins possible theoretically is 20^{299}.

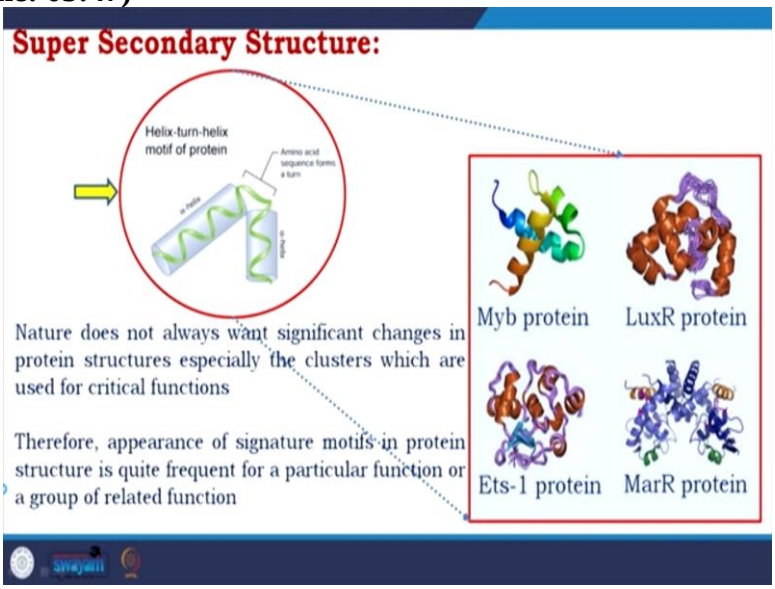
You would have to multiply the number of possible amino acids in each position by the next number of possible amino acids using the multiplication rule. So the theoretical number of combinations would be 20^{299}.

(Refer Slide Time: 04:43)



You will be surprised to see that the number is that big. But this number of proteins does not generate in reality. We call natural intelligence the creativity, the smartness nature poses to optimize things. So, one reason we will discuss in the next class is the folding possibilities. But there are other reasons, and those show the efficiency of NI. Natural intelligence (NI) and I would introduce super secondary structure in this context.

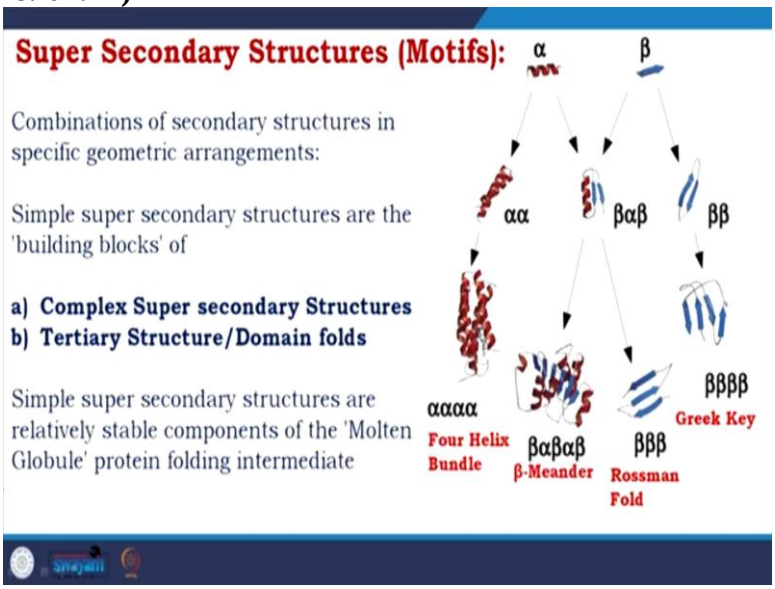
(Refer Slide Time: 05:47)



Nature does not always want significant changes in protein structures, especially in the clusters used for critical function. So think about the function of nature. Every organism has its way, and every way needs proteins. So, many proteins are needed, and natures have to develop the rule of their folding. But nature being creative, smart, and wonderful, managers try to find the commonality.

So, Mother Nature has created some clusters, some chain of amino acids, some specific folds which are determined or destined to do some specific function. So therefore, the appearance of a signature motif in protein structure is quite frequent for a particular function or a group of related functions. How wonderful? Taking an example, Helix turn helix is a very common motif. And you see that in Myb protein, in LuxR protein, in Ets 1 protein, in MarR protein, everywhere you could see the presence of helix turn helix. So that is what natures intelligence is. She puts one motif in many structurally divergent proteins but has one common function called super secondary structure or motif.

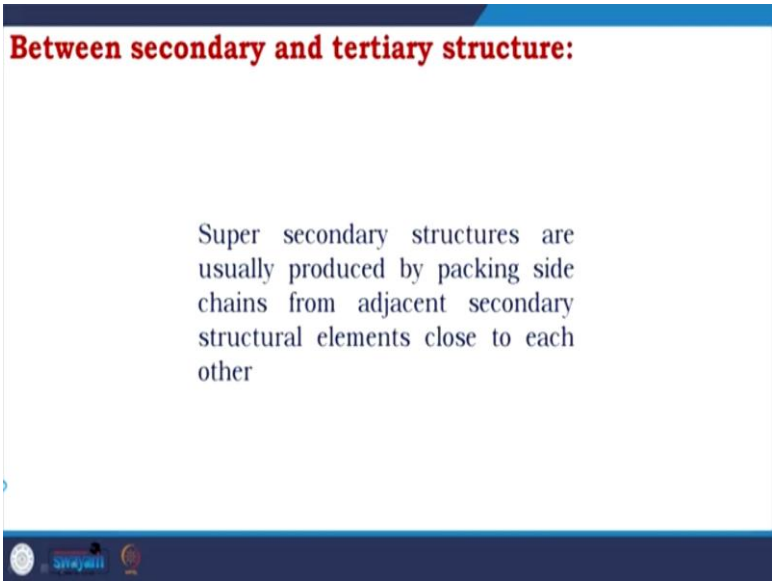
(Refer Slide Time: 07:42)



What are super secondary structures? They are the combination of secondary structures in specific geometric arrangements. You have alpha helix. You have a beta sheet. So alpha-alpha, beta-beta, alpha-beta, this type of permutation and combination develops super secondary structures. Super secondary structures are the building blocks of complex super secondary structures and tertiary structures or domain folds.

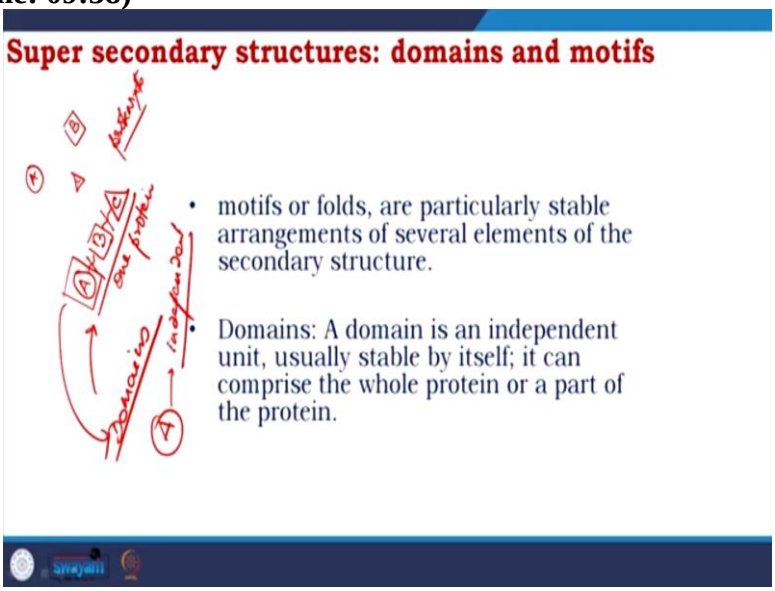
Simple, super secondary structures are relatively stable components of the molten globule of protein folding intermediate. The example of alpha-alpha-alpha-alpha four-helix bundle, beta-alpha-beta-alpha-beta is beta Meander, beta-beta-beta is Rossman fold, and beta-beta-beta-beta is Greek key motifs.

(Refer Slide Time: 09:21)



So, super secondary structures are usually produced by packing side chains from adjacent secondary structural elements close to each other. And as it floats, it is between the secondary and tertiary structure.

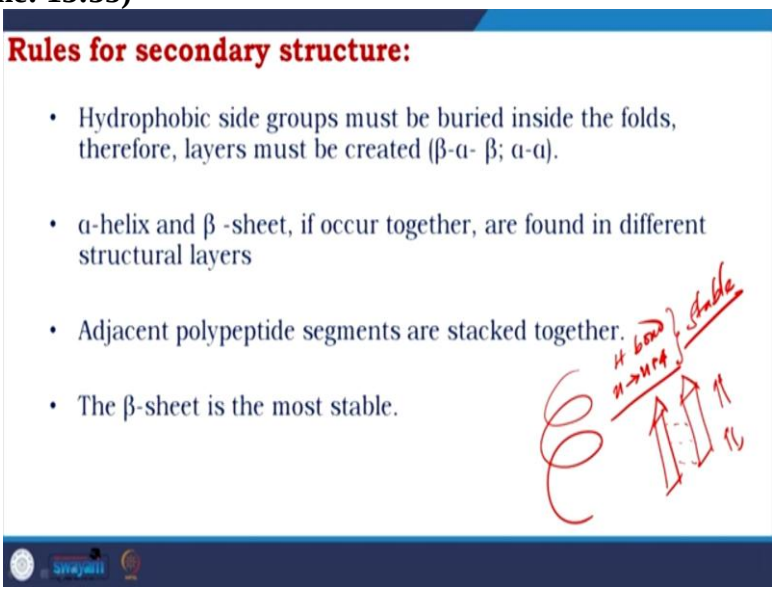
(Refer Slide Time: 09:38)



Domain and motifs are some of the most confusing terms to separate. Motifs or folds are particularly stable arrangements of several secondary structure elements. A domain is an independent unit usually stable by itself; it can comprise the whole protein or a part of the protein. You have to experiment with the best way to separate the domains and motifs. If you have a protein domain and a motif, you design them, synthesize them separately, try to produce that protein, purify that protein. If you could get the protein produced and purified, it is a domain. Otherwise, it is a motif. So as I told domains or motifs, talk about specific functions. But domains have their independent existence.

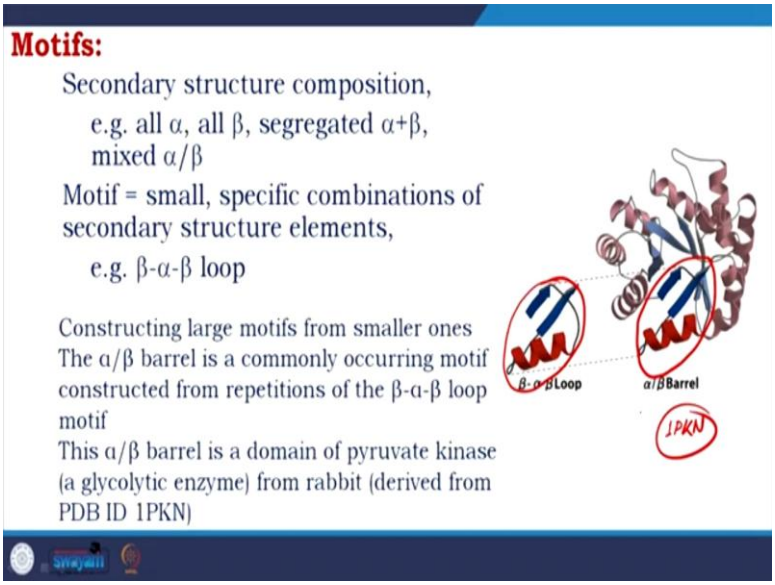
Do you understand what a domain is if you think about evolutionary theory? There is evidence where it was understood or found in the literature that, if there is much protein, let us say protein A B C, in a prokaryotic system, in the course of evolution, they come into a more complicated eukaryotic organism, where the function has to be done with much faster speed because in a complex organism there are more number of protein hierarchy of the function and all these things. In evolution, they come under one protein. In this situation, each of them is called a domain. Motifs are part of the protein or domains, but they cannot work independently.

(Refer Slide Time: 13:35)



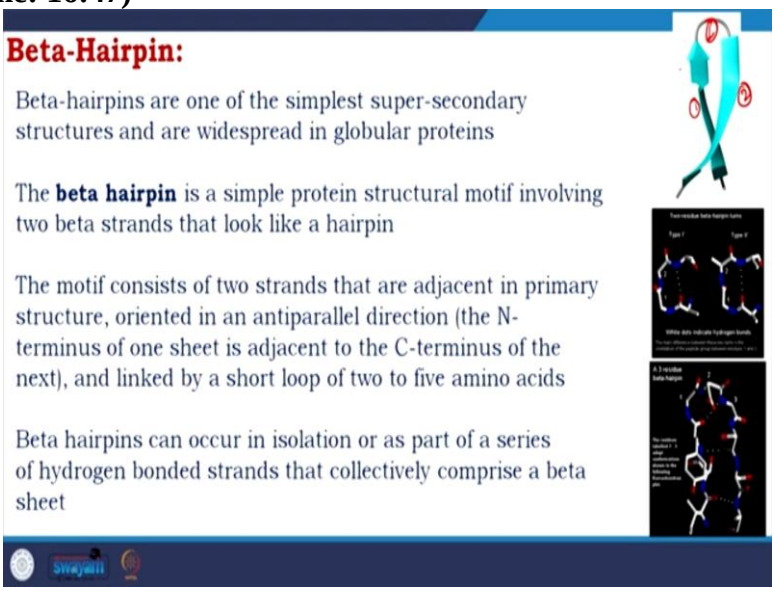
Rules for secondary structure development: Hydrophobic side groups must be buried inside the folds; therefore, there will be layers if you do that. So beta-alpha, beta-alpha-alpha layers are created. Alpha helix and beta sheets are found in different structural layers. Adjacent polypeptide segments are stacked together, and the beta-sheets are most stable. Beta strands cannot stay independently, so they are together. Together they have hydrogen bonding if they are parallel or anti-parallel. So when they are together, they bring more stability more energy to the system.

(Refer Slide Time: 15:24)



Secondary structural composition, all alpha, all beta, segregated alpha + beta, mix alpha-beta, the motif is a small specific combination of secondary structure elements connected by loops. Constructing large motifs from smaller ones, the alpha beta-barrel is a commonly occurring motif constructed from the repetition of the beta-alpha-beta loop motif.

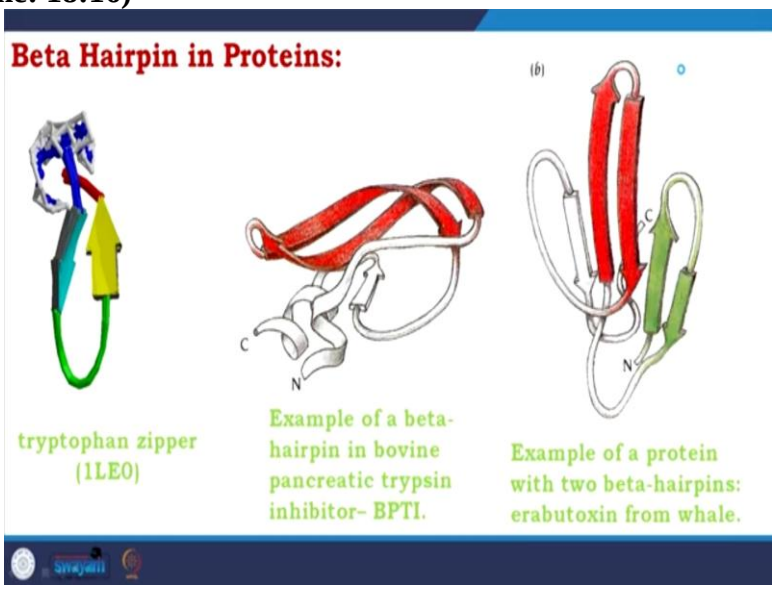
(Refer Slide Time: 16:47)



Beta-hairpin is a very known motif. One of the simplest super secondary structures is widespread spatially in globular proteins. The beta-hairpin is a simple protein structural motif involving two beta-strands that look like a hairpin. As you see here, one beta-strand two beta strands connected with a loop and looked like a hairpin. The motif consists of two adjacent strands in a primary structure oriented in an anti-parallel direction. The N terminus of one sheet is adjusted to the C terminus of the next and linked by a short loop of two to five amino acids. So depending on how many amino acids develop loop, they have classifications like

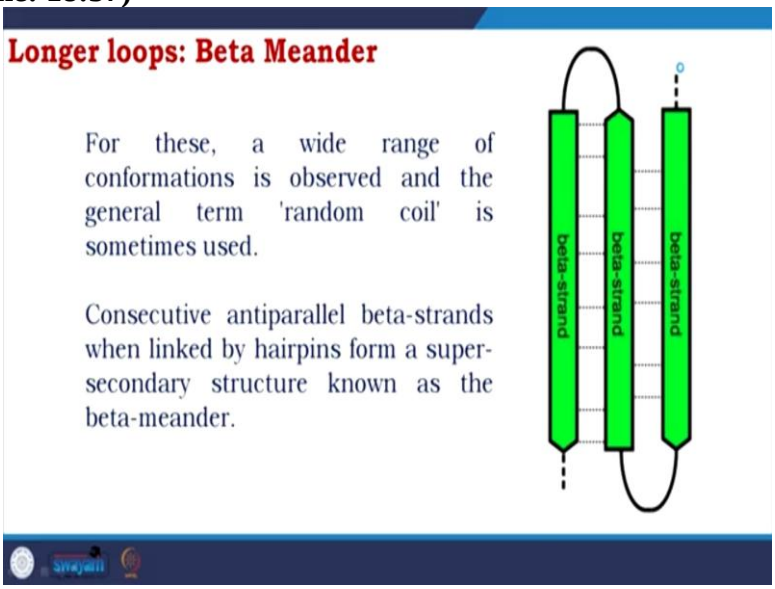
two residue beta-hairpin and three residue beta-hairpin. Beta-hairpins can occur in isolation or as a part of a series of hydrogen-bonded strands that collectively comprise a beta-sheet.

(Refer Slide Time: 18:16)



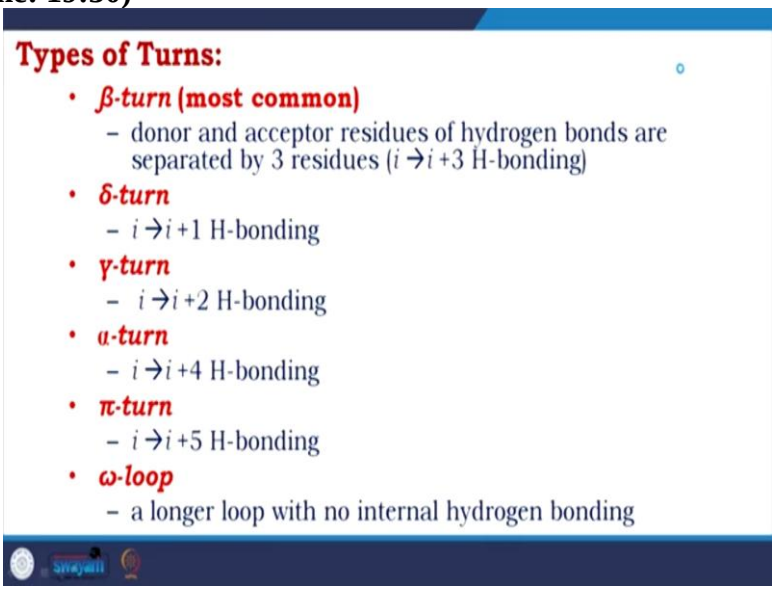
An example of beta hairpins presence in protein is the tryptophan zipper. This is having the beta-hairpin you could see. This is an example of beta-hairpin in bovine pancreatic trypsin inhibitor BPTI. BPTI fold is a classic stable and frequently found fold in nature, an example of a protein with two beta hairpins. Erabutoxin from the whale, beta-hairpin, also has an interesting aspect: their presence is in many toxin proteins.

(Refer Slide Time: 18:57)



Coming to another one with longer loops is called beta meander. A wide range of conformation is observed for these. The general term random coil is sometimes used—consecutive anti-parallel beta strands when linked by hairpins from a super secondary structure known as a beta meander.

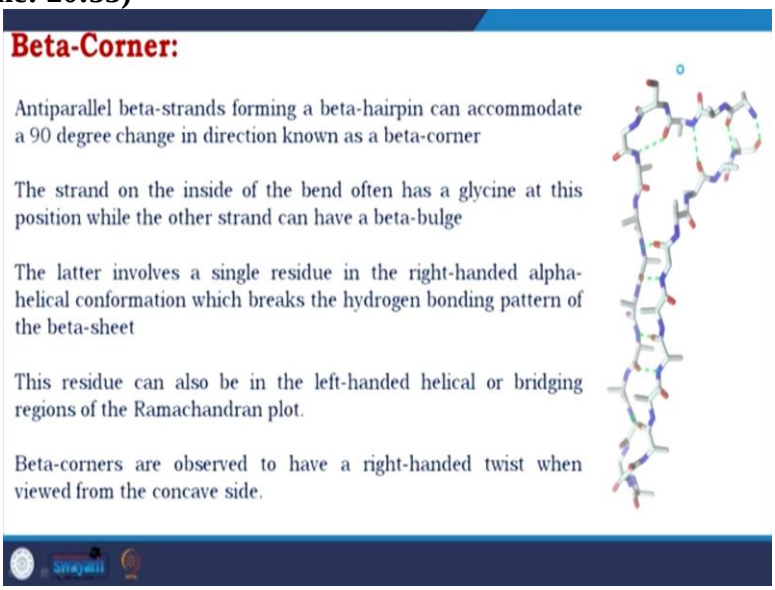
(Refer Slide Time: 19:30)



I want to talk about different turns and loops. We have already talked about beta turns, donor and acceptor residues of hydrogen bonds are separated by 3 residues. So i to i + 3 hydrogen bonding, but there are many other turns like delta turn, which is i to i + 1 hydrogen bonding, gamma turn, which is i to i + 2 hydrogen bonding, alpha turn, which is i to i + 4 hydrogen bonding, pi turn which is i to i + 5 hydrogen bonding and omega loop, a longer loop with no internal hydrogen bonding.

Omega loop is present in many proteins, and when present, they take a very interesting role in catalysis, they influence the active side volume, they interact with a substrate, one of the very popular examples of omega loop is class A beta-lactamase.

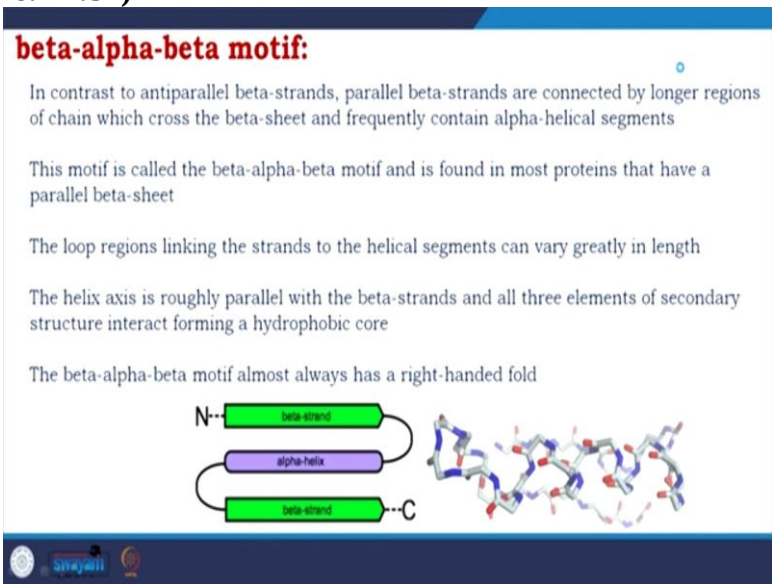
(Refer Slide Time: 20:53)



Anti-parallel beta-strands forming a beta-hairpin can accommodate a 90-degree change in the direction known as a beta-corner. The strand on the inside of the bend often has a glycine at

this position so that the flexibility comes while the other strand can have a beta-bulge, the extension. The latter involves a single residue with the right-handed alpha-helical conformation, which breaks the beta-sheet's hydrogen-bonding pattern and helps it turn. This residue is also in the left-handed helical or bridging regions of the Ramachandran plot, as discussed in our last class. When viewed from the concave side, beta corners are observed to have a right-handed twist.

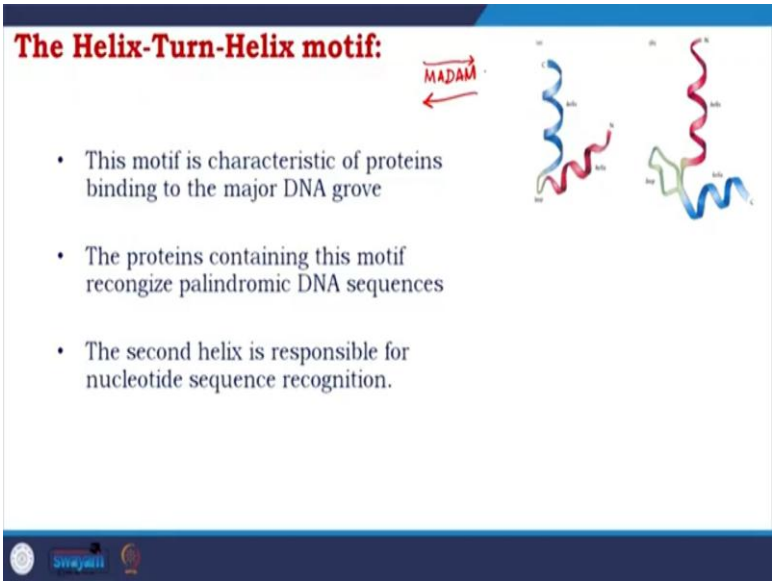
(Refer Slide Time: 21:52)



Beta-alpha-beta motif, so up to now, the motifs we talked about concerning beta-strands are mostly anti-parallel. But beta-alpha-beta motif is where parallel beta-strands are connected by a longer chain region that crosses the beta-sheet and frequently contains alpha-helical segment. So it is beta-beta in between alpha. The motif is called the beta-alpha-beta motif, and it is found in most proteins with a parallel beta-sheet.

The Loop regions linking the strands to the helical segments can vary greatly in length. The helix axis is roughly parallel with the beta-strands, and all three elements of the secondary structure interact, forming a hydrophobic core. The beta-alpha-beta motif almost always has a right-handed fold.

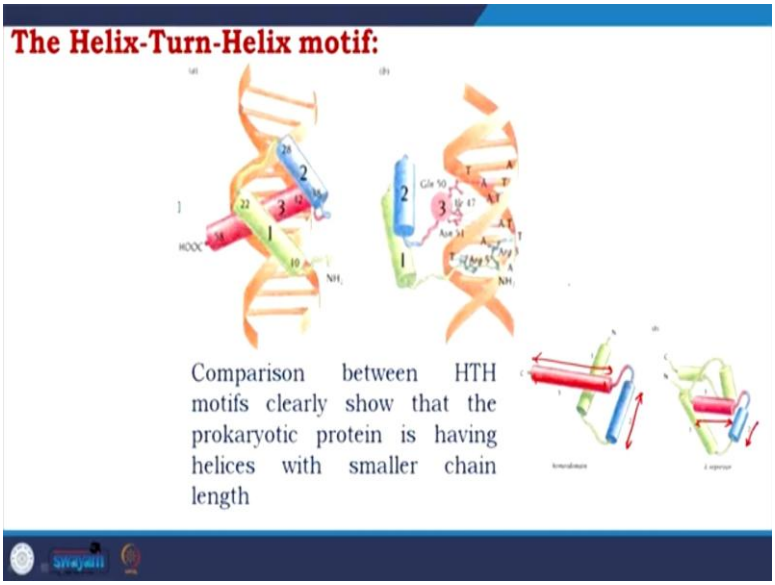
(Refer Slide Time: 22:51)



The helix-turn-helix motif started our discussion of a super secondary structure. It is a very popular one. It is characteristic of protein binding to the major DNA group. So if you know about DNA, in DNA, there is a major grove and a minor grove. So helix-turn-helix, it goes there and binds to the major group. The proteins containing this motif recognize the palindromic DNA sequence. Palindromic is, if you have a word that could read from both sides coming up with the same thing, it is called palindromic.

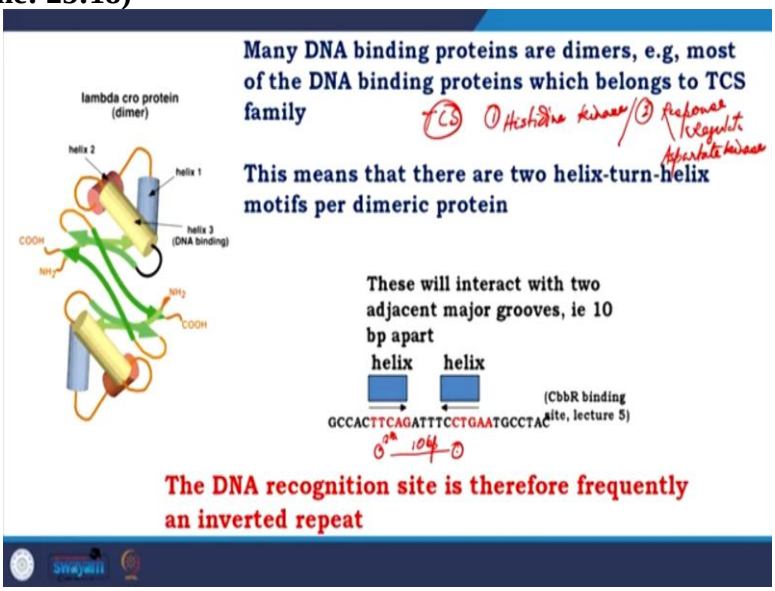
Similarly, you start reading from both sides of a DNA sequence and see or get the same sequence. The second helix is responsible for nucleotide sequence recognition. So the first alpha-helix goes and binds to the major group and the second helix is responsible for nucleotide sequence recognition.

(Refer Slide Time: 24:21)



So you see the binding in the helix-turn-helix motif. Very interestingly, here, you have two helix-turn-helix motifs. One is the homeodomain from the eukaryote, and one is the lambda repressor in the prokaryote. So it clearly says that, if you compare, the prokaryotic protein has helices with smaller chain lengths. So that says that there is a change in evolution from prokaryotic to eukaryotic. The change is the sizes of the helices are significantly larger in higher eukaryotes.

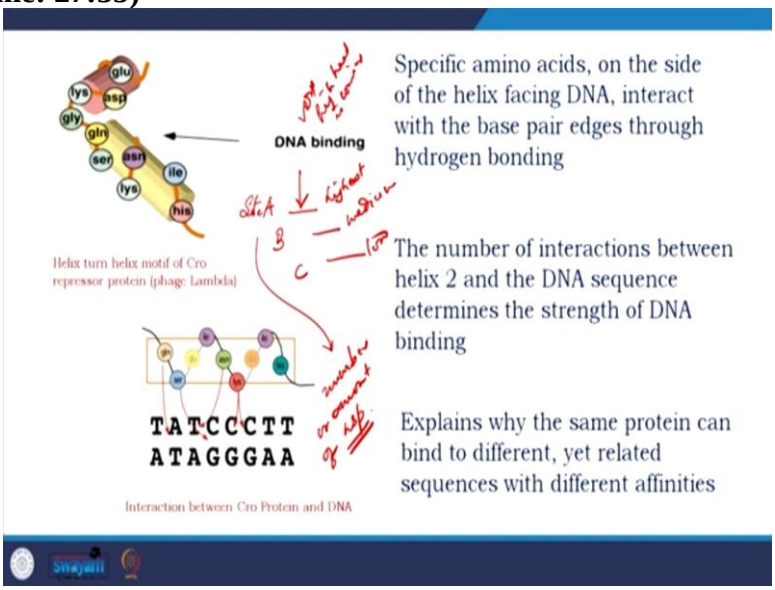
(Refer Slide Time: 25:18)



Many DNA binding proteins are dimer. For example, most DNA binding protein belongs to TCS (two components system family). The two-component system is a system that we will discuss later. It is a very interesting system that helps bacteria survive in their critical moment. There is one Histidine kinase component in the TCS system, and two response regulators called Aspartate kinase.

So they bind by making dimers. This means that there are two helix turn helix motifs per dimeric protein, one from monomer A and the other from monomer B. This will interact with two adjacent major groups that are 10 base pairs apart. Nowadays, it is similar patterns used in technology development. So this is an amazing pattern.

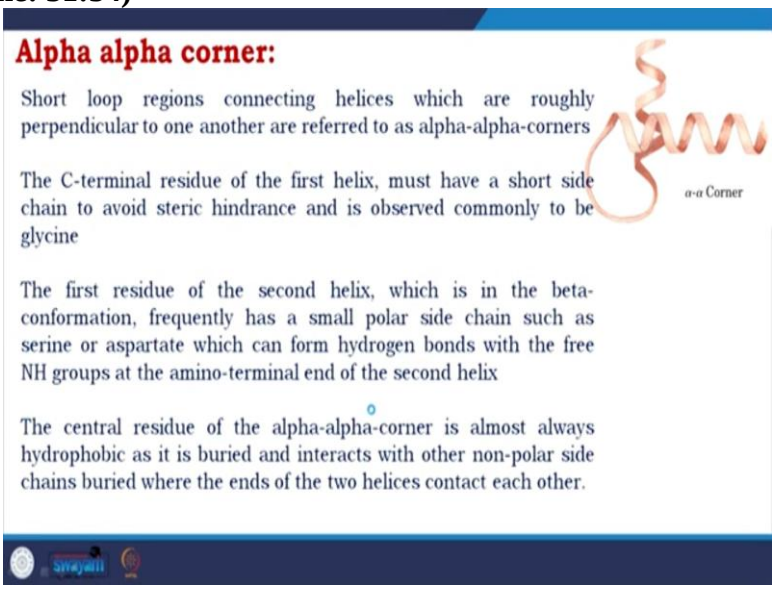
(Refer Slide Time: 27:35)



Looking at the DNA binding closer, you will see that specific amino acid on the side of the helix facing DNA interacts with the base pair edges through hydrogen bonding. So there is phosphate, phosphates are taking part in the interaction, but the nucleosides are also. The number of interactions between helix two and the DNA sequence determines the strength of DNA binding. So the first helix goes to the major group, the second helix interacts with the nucleotides, and depending on the difference in the nucleotide, the interaction trend differs. So what, what is the significance? The significance is I told nature is extremely smart. So these proteins, these DNA binding proteins, tell the organism about a situation. Let us say it tells the bacteria that somehow, where you are staying, the heat is increasing, so you have to escape if you want to survive. Now, heat is not a specific term. It is a relative term. It might be a low amount of heat, medium amount of heat, high amount of heat, and even many other things in between. How could bacteria sense by getting the signal that this is very high heat or a low amount of heat? That is, binding 2 nucleotides do the difference in different proteins. So it comes, and it binds to a specific site to get interaction. It is the same protein, but binding to this site, it would provide a message that very high heat is coming. Here medium heat is coming. Here the low amount of heat is coming. Why it is very critical, depending on them, this will produce the number or amount of heat shock proteins to help bacteria survive.

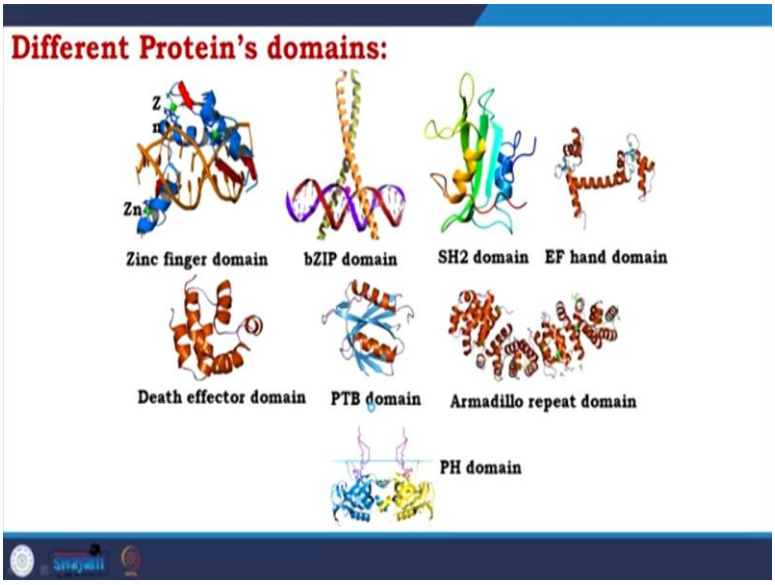
So this explains why the same protein can bind to different yet related sequences with different affinities giving different signals.

(Refer Slide Time: 31:34)



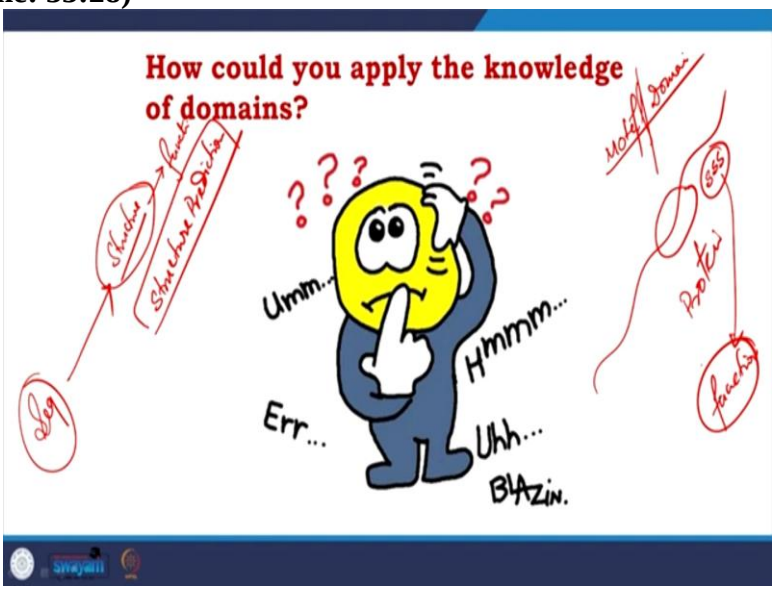
Coming to another alpha-alpha corner, short loop region connecting helices that are roughly perpendicular to one another are referred to as alpha-alpha corner. You see the picture of the alpha-alpha corner, two alpha helices connected by a loop. The C terminal residue of the first helix must be a short side chain to avoid steric hindrance and is observed commonly to be glycine. The first residue of the second helix in the beta conformation frequently has a small polar side chain such as serine or aspartate, forming a hydrogen bond with the free NH group at the amino-terminal end of the second helix. The alpha-alpha corner's central residues are almost always hydrophobic as it is buried and interacts with other non polar side chains where the ends of the two helices contact each other.

(Refer Slide Time: 32:49)



The zinc finger domain is also a DNA binding domain, bZIP is another DNA binding domain. I am generally talking about DNA binding domains because DNA binding domains are mostly transcription factors, work with signal transfer, and are very critical and very critical well understood mostly. SH2 domain is not DNA binding. EF-hand domain, Death effector domain, PTB domain, Armadillo repeat domain, and PH domain. They are some commonly known domains.

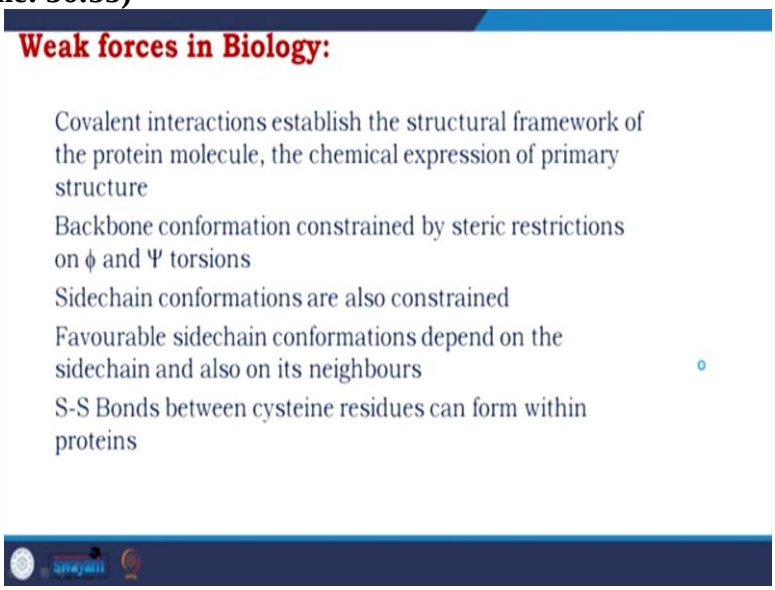
(Refer Slide Time: 33:28)



So how could you apply the knowledge of domains? While discussing, I have already introduced the importance. I want to end up with that super secondary structural motif or domain with another critical role. When I introduced the course, I am repeatedly talking about one of the biggest, most critical problems I talked about, sequence we get sequence from NGS.

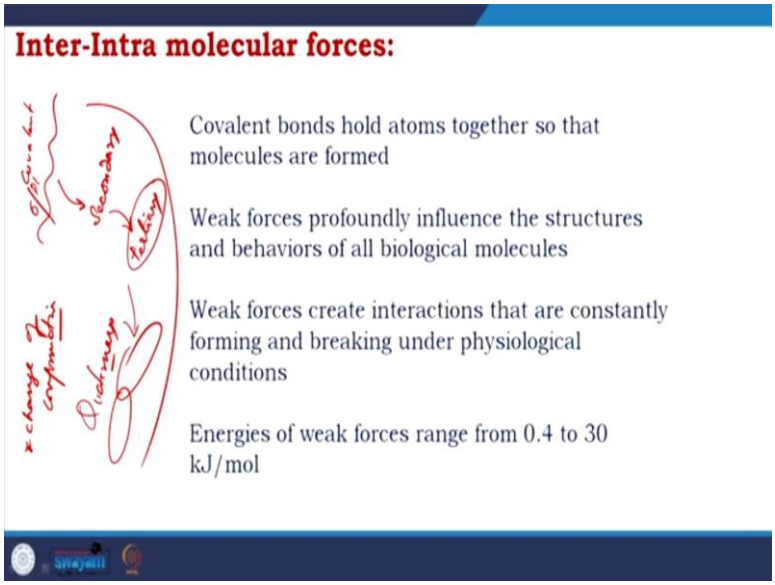
We want to get function, and we need structure. We will continue with the structural biology technique and details in the next modules. Still, in our previous introductory module, I have already discussed it is impossible in the present scenario to make a huge speed up in structure development techniques to connect the bridge. The alternatives and one of the most accepted alternatives is structure prediction. When I am talking about this, I know some of you definitely argue about the artificial intelligence about alpha fold, which is impressively solving structure problems now. Still, if you look at the bigger picture, unfortunately, there are many more things to do. So when we are doing structure prediction, be it any prediction modeling method, simulation energy minimization, or artificial intelligence, getting the knowledge of domains would be immensely helpful in understanding a protein.

(Refer Slide Time: 36:55)



Coming to the next part of the topic today, weak forces. Weak forces in biology covalent interaction established the protein molecule's structural framework, the primary structure's chemical expression. Backbone conformation constrained by steric restrictions on φ and ψ torsion angles. Sidechain conformations, determined by the χ torsion angle, are also constant. Favorable sidechain conformation depends on the side chain and also on its neighbors. So it is not only about the amino acid present there. It is also about the neighbor. If you have glycine and glycine, if you have glycine and arginine, if you have arginine and arginine, they will behave very differently. And about disulfide bonds, they are the only covenant bond that could form in the physiological condition they could alter.

(Refer Slide Time: 37:56)

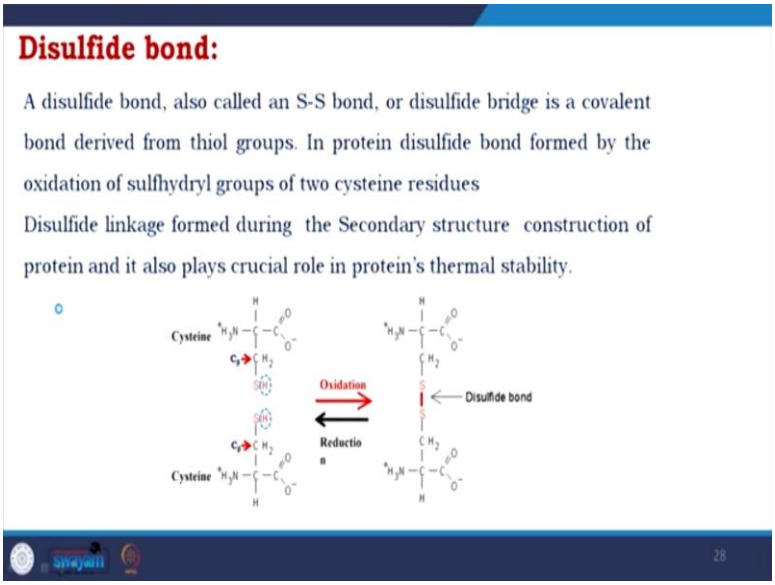


Inter and intramolecular force, the covalent bonds hold atoms together to form molecules. Weak forces profoundly influence the structures and behaviors of all biological molecules. Weak forces create interactions that are constantly forming and breaking under physiological conditions. The energies range from 0.4 to 30 kj/mol.

So what we are talking about? We are talking about that a protein primary structure we get from covalent bonds, mostly sigma, but you also have pi. Then our journey is from this to secondary then tertiary and this ends with one, if you have more than one then quaternary. This part; the journey there and also the change of conformation, why change of conformation is very important, because for protein molecule to be functional it have to show the change.

When it want to start a reaction, when you want to finish the reaction, when it wants to work, when it do not want to work. All these happen through the change of conformation show change of conformation is very important this needs breaking and making up non covalent bonds.

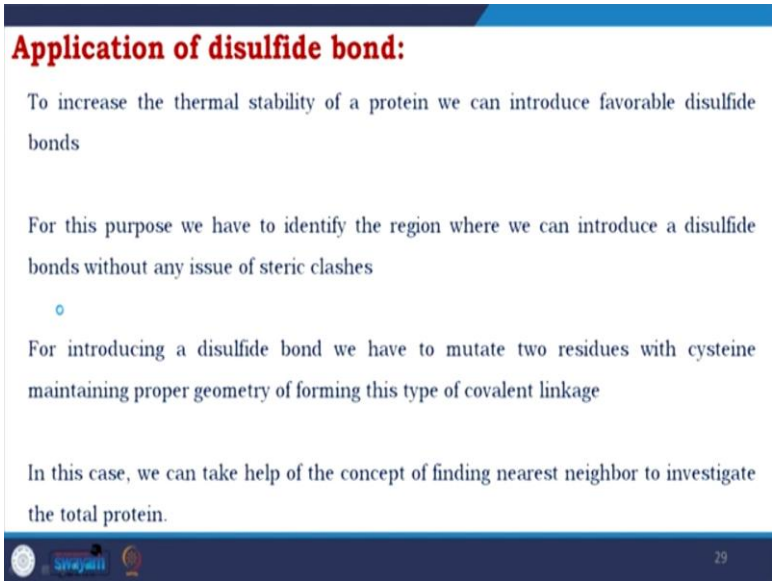
(Refer Slide Time: 39:43)



So before going into non covalent as I told I will discuss about disulfide bond, a disulfide bond also called a S S sulphur sulphur bond or di sulfide bridge is a covalent bond derived from thiol groups. In protein disulfide bond formed by the oxidation of sulfhydryl group of two cysteine residues. We have only two amino acids, cysteine and methionine which have sulfur, but methionine having a CH three (()) (40:10) group it cannot be taken out.

So it could not form disulfide, cysteine is the only amino acid forming disulfide bond. Disulfide linkage formed during the secondary structure construction of protein and it also plays a crucial role in proteins thermal stability by introduction of one or more disulfide bond you could increase considerably thermal stability of a protein. So you see, when two cysteine they are in close proximity in the protein structure in oxidizing condition they form disulfide bond.

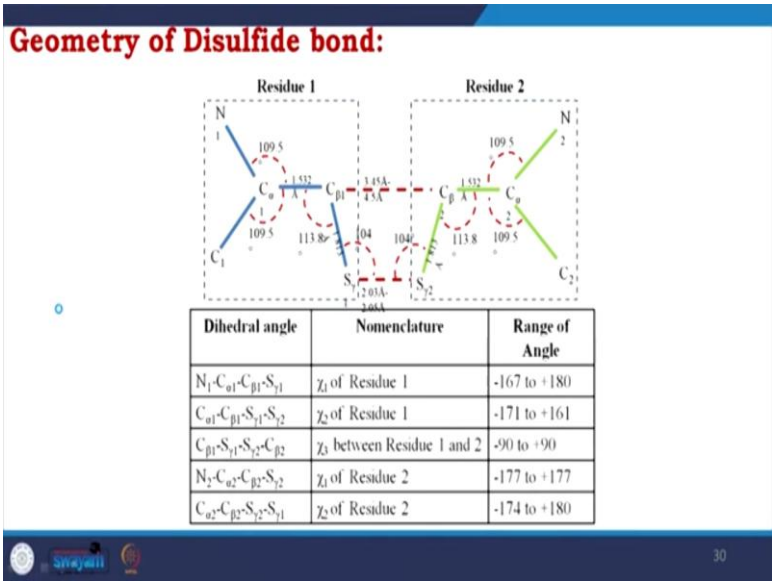
(Refer Slide Time: 40:48)



Application of disulfide bond to increase the thermal stability of protein we can introduce favorable disulfide bonds. So, when we compare between two proteins, the presence of disulfide bond differs, but this gives us idea about engineering. When you have a protein you want to make it thermo stable you could have approximated or strategized introduction two sulfur or two cysteine, so that they could form a disulfide bond.

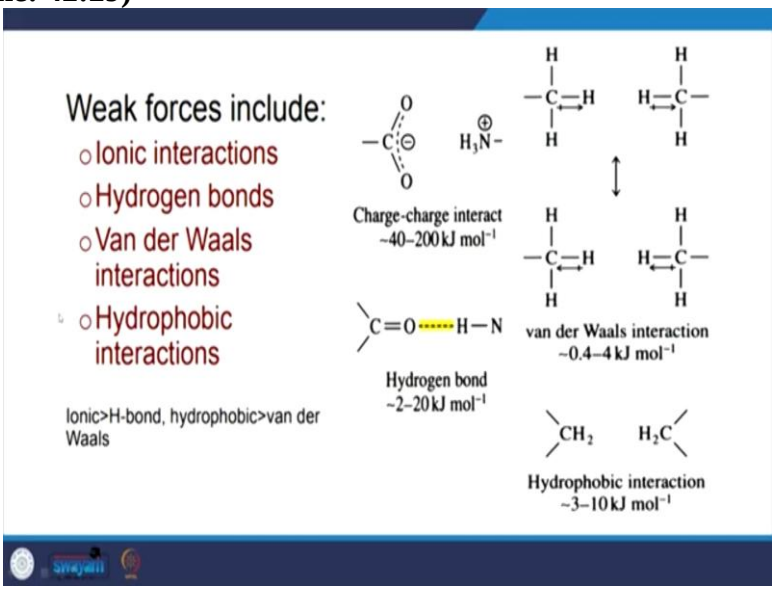
For this purpose we have to identify the region where you can introduce a disulfide bond without any issue of steric clashes. For introducing a disulfide bond we have to mutate two residues with cysteine, maintaining proper geometry of forming this type of covalent linkage. In this case we can take help of concept of finding nearest neighbor to investigate the total protein. We will talk about finding nearest neighbor concept when we will discuss the enzyme engineering or protein engineering. So when you have to do this calculation, you have to understand the geometry very well.

(Refer Slide Time: 42:07)



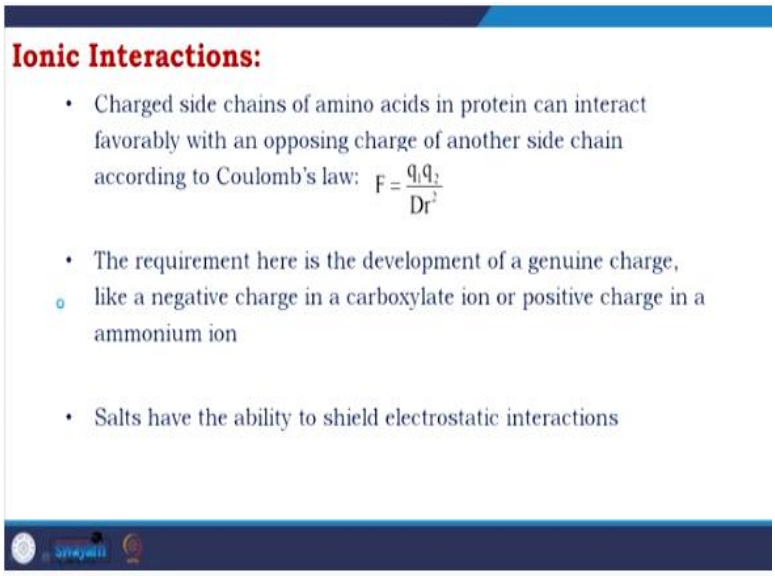
So I have given the details of 2 cysteine collected from many proteins who are successfully forming disulfide bond. This information being into your database would help you to design a new cysteine pair in a protein and we will discuss about that in details.

(Refer Slide Time: 42:29)



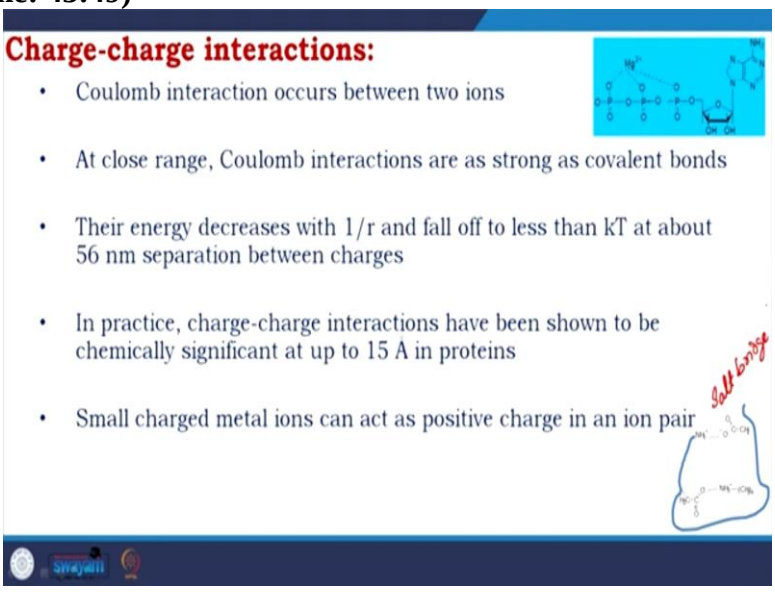
So weak forms include ionic interaction, hydrogen bond, Van der Waal interaction and hydrophobic interactions, there are base tracking and all these things but mostly these four are taken care. Ionic interaction is a real charge-charge interaction; the presence of real charge makes it happen in energy close to covalent bond. Hydrogen bond as approximated 2 to 20, it differs depending on the situation but this is a low energy. Van der Waal interaction, 0.4 to 4 kilojoules per mole and Hydrophobic interaction 3 to 10 kilojoules per mole depending on the patches.

(Refer Slide Time: 43:14)



So ionic interaction, as I told charge side chains of amino acids in protein can interact favorably with opposing charge of another side chain according to Coulombs law, F = q 1 q 2 by Dr 2, where q 1 and q 2 are two charge entity, r is the distance, d is that dielectric constant. The requirement here is the development of genuine charge like a negative charge in a carboxylate ion or positive charge in it ammonium ion. Salts have the ability to shield the electrostatic interaction.

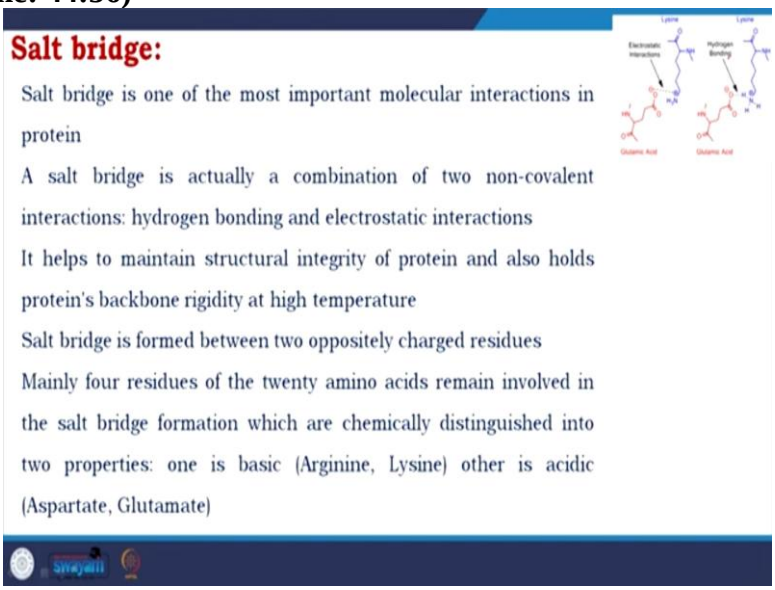
(Refer Slide Time: 43:49)



So charge-charge interaction, it happened in protein as I told but it could happen between metal with the phosphate in nucleoside or any phosphorylated protein. So coulombic interaction occurs between two ions in a simple term at close range coulomb interactions are as strong as covalent bonds, their energy decreases with 1 / r. So with increasing distance, energy decreases and follows up to less than kT thermal energy at about 56 nanometers separation between charges.

In practice, charge-charge interactions have been shown to be chemically significant up to 15 angstrom in proteins. Small charged metal ions can act as positive charge in an ion pair. So these are another example of amino acids. So when this happens, this ionic interaction you see here, we call it salt bridge. So let us talk about salt bridges.

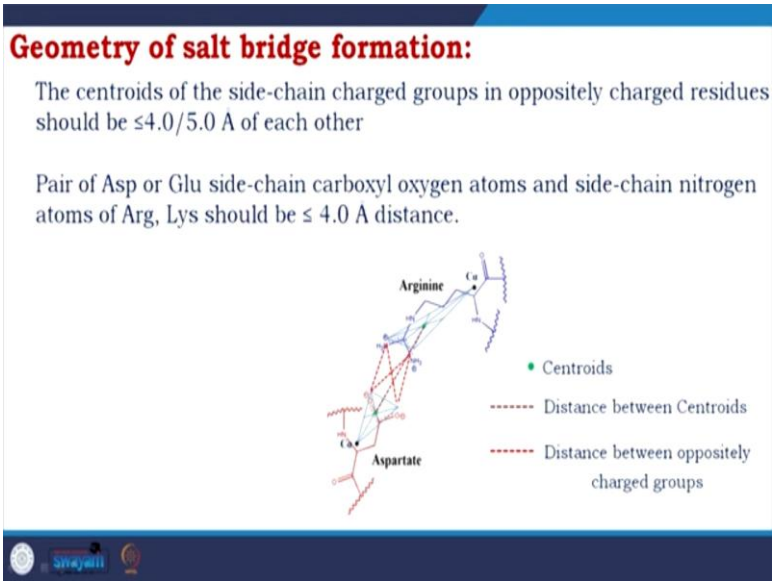
(Refer Slide Time: 44:56)



The salt bridge is one of the most important molecular interactions in protein. A salt bridge is actually a combination of two non covalent interactions, hydrogen bonding and electrostatic interaction. If you see here between glutamic acid and lysine, you see that this is the charge, real charge form and they are having the electrostatic interaction but when it is not between the real charges the hydrogen comes inside, it is called hydrogen bond.

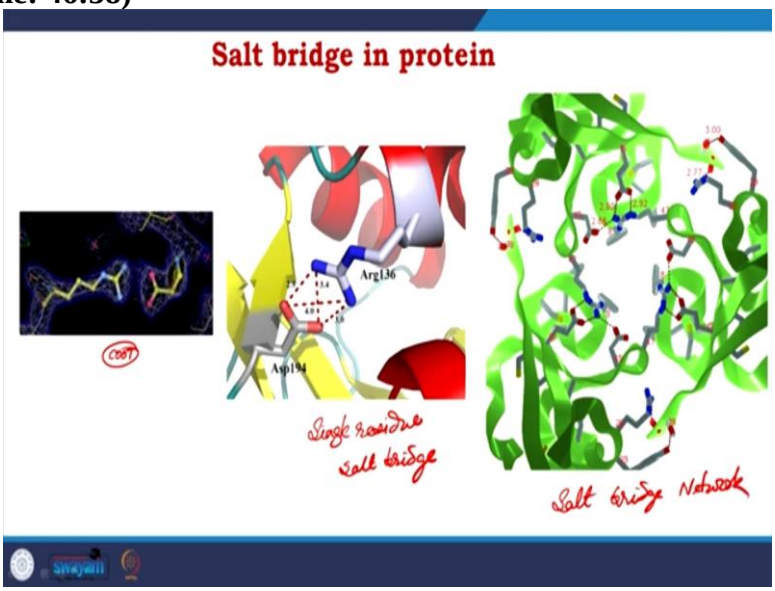
It is very difficult to separate them conceptually. It helps to maintain structural integrity of a protein and also holds proteins backbone rigidity at high temperature. Salt bridge is formed between two oppositely charged residues. Mainly four residues of 20, amino acid remain involved in salt bridge formations which are chemically distinguished into two properties. One is basic which is arginine and lysine, as others are acidic, aspartate and glutamate.

(Refer Slide Time: 45:58)



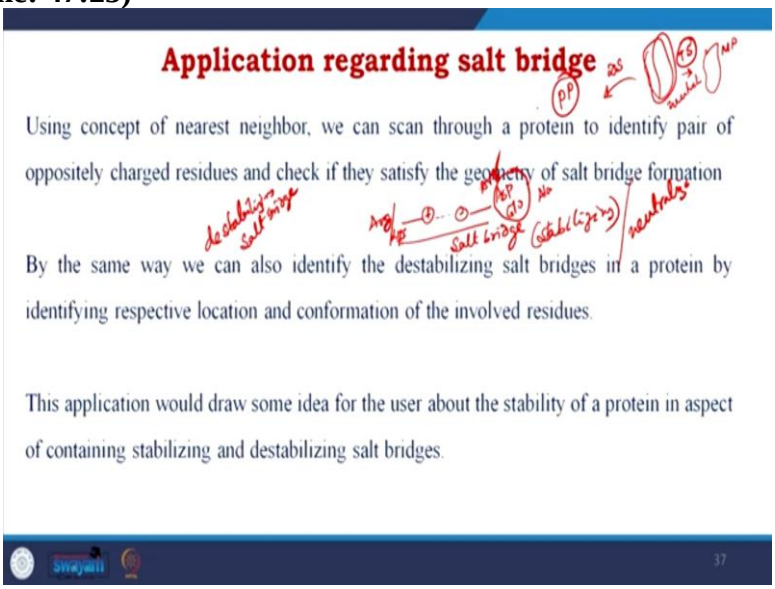
So geometry of salt bridge is very important. The centroids of the side chain charged group in oppositely charged residues should be less than equal to 4 to 5 angstrom from each other. Pair of aspartate or glutamate side chain carboxyl oxygen atoms and side chain nitrogen atoms of arginine and lysine should be less than equal to 4 angstroms. So here if you see the green spot, it talks about the centroid. The brown dotted lines are the distance between centroids and the red dotted lines are distance between oppositely charged groups.

(Refer Slide Time: 46:38)



So these are different example of solid region protein. This is presented in a crystal structure where you see the electron density. This is presented in a, by a software coot which we will discuss. Then this is a glimpse where one arginine and one aspartate are developing and this is a salt bridge network. So this is a single salt bridge, single residue salt bridge, this is salt bridge network. So you see the differences.

(Refer Slide Time: 47:25)



Application using concept of nearest neighbor which I talk in case of disulfide, we can scan through a protein to identify pair of oppositely charged residues and check if they satisfy the geometry of the salt bridge formation. By the same way we can also identify that destabilizing salt bridges. So destabilizing salt bridge is the concept of which I really worked on personally, this is a amazing concept of nature.

Suppose this is a salt bridge, because it is oppositely charged we are telling it as a stabilizing salt bridge. Now a stabilizing salt bridge would help stabilize a protein. So here you know that positive charge you have arginine or lysine and here you have aspartate or glutamate. Now you see, if you muted this let us say to alanine then the positive effect would be neutralize they would no more stabilize.

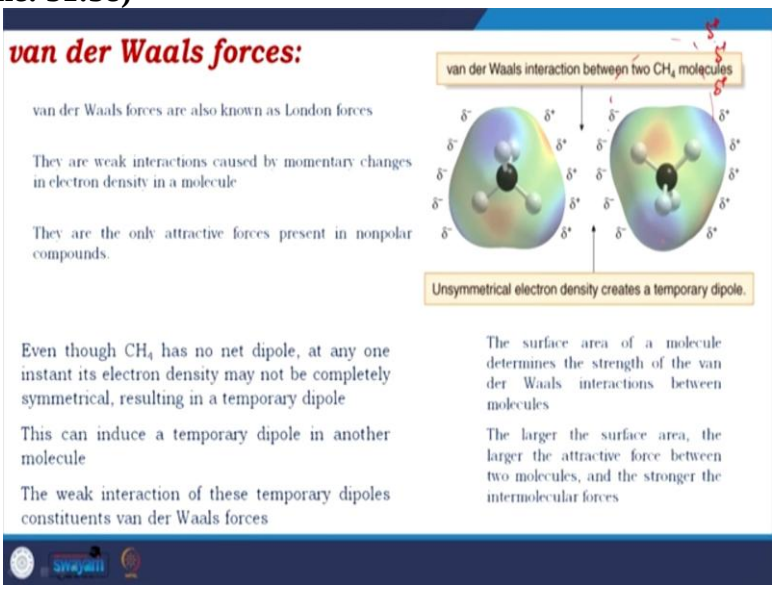
Think about instead of doing this, if you model a same charge like arginine here let us say then they will repel each other. So that is what we say destabilizing salt bridge. Now if you are listening carefully and understanding the concept, you will say why nature will do that you say nature is smartest. And by making a destabilizing salt bridge, nature would actually introduce adverse effect in a protein.

But if you think what I am saying, suppose you have a thermo stable protein nature want to make it mesophilic protein. So nature did that neutralizing effect. But if it do that destabilizing salt bridge then from this thermo stable it could make a protein which work in cold conditions that is called psychrophilic protein. So you understand what I am saying. So

when you have a thermophilic protein by introducing a destabilizing salt bridge concept, you introduce huge flexibility on that portion.

And in course of evolution that might be performed to convert a thermostable protein to a psychrophilic protein. That is where we introduce the concept of destabilizing, by identifying the destabilizing amino acid or destabilizing salt bridge and change it to a stabilizing salt bridge would have double effect towards stabilization. So this is a very interesting strategy in thermo stable enzyme or thermo stable protein engineering. The application would draw some idea for the user about the stability of a protein in aspect of containing stabilizing and destabilizing salt bridges.

(Refer Slide Time: 51:38)



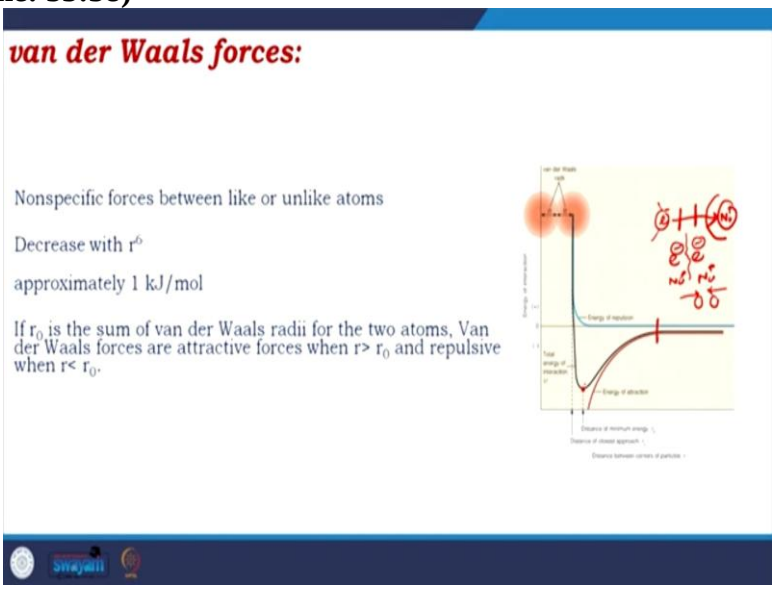
Coming to the next topic van der Waal forces, van der Waal forces are also known as London forces. They are weak interaction caused by momentary changes in electron density in a molecule. So in one word, if you have a molecule and because of its stretching, bending somehow, because of dynamic changes, the electron distribution changes, there is a dipole development that is where you get the van der Waal.

They are the only attractive forces present in non polar compounds. As an example, even though methane has no net dipole at any one instant its electron density may not be completely symmetrical, resulting in a temporary dipole. So as I was discussing about as this picture is discussing about, so when there is bone stretching and all, there is difference in the electron density distribution.

As you see here unsymmetrical electron density that creates a dipole. So you see here a presence of partial negative and partial positive charges. This can introduce a temporary dipole in another molecule that weak interaction of these temporary dipoles constituents of Van der Waals forces. The surface area of a molecule determines the strength of the van der Waals interaction between molecules.

The larger the surface area, the larger the attractive force between two molecules and the stronger the intra molecular. So if you see there is some amount of negative and positive charge develop. Now if the area would be doubled then you have more amount of charge difference.

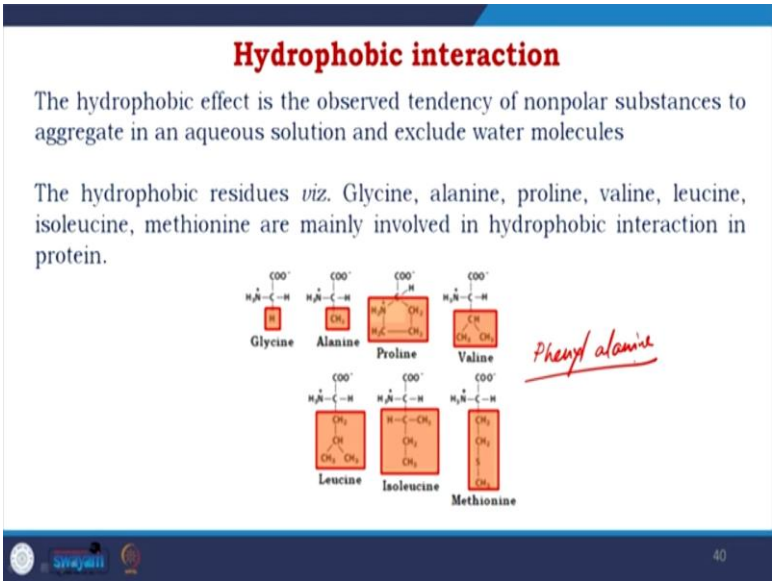
(Refer Slide Time: 53:38)



So how it happened if you have two atoms coming together initially the electrons of one attracts the nuclei of other. So when it starts from this point, because of this attractive force, there is energy reduced. So this is energy of attraction between the electrons and nuclei. But when it comes to certain distance then electron, electron, nuclei, nuclei repulsion started and it would increase in the energy.

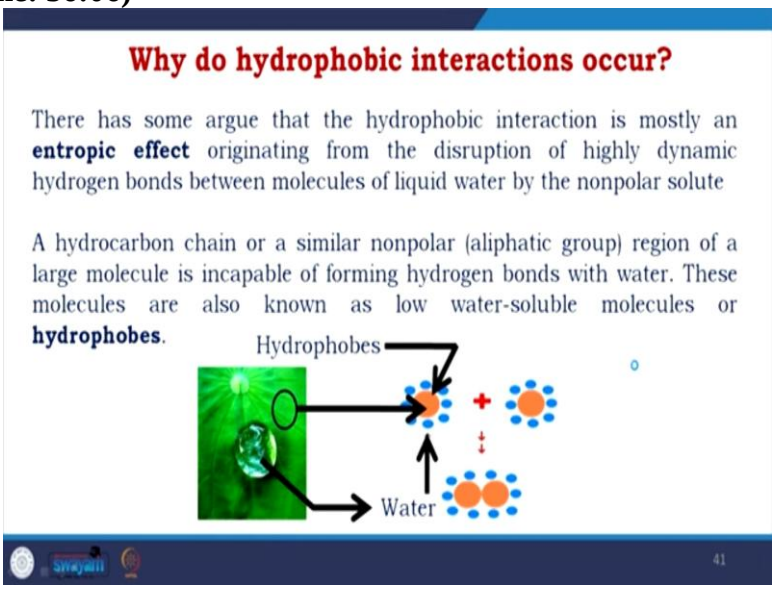
This is the minimum energy distance and this is the equilibrium distance it decrease with the power of r to the power 6, where r is the distance approximately one kilojoules per mole. If r 0 which is the equilibrium distance is the sum of van der Waal radii of the two atoms, van der Waals forces are attractive forces in this region when r greater than r 0 and repulsive when r less than r 0 in this, this is the repulsive, this is the attractive.

(Refer Slide Time: 55:20)



Coming to the next one hydrophobic interaction, the hydrophobic effect is the observed tendency of non polar substances to aggregate in an aqueous solution and exclude water molecule. So they are hydrophobic. So they fear of water they want to exclude the water. The hydrophobic residues glycine, alanine, proline, valine, leucine, isoleucine, methionine are mainly involved in hydrophobic interaction in protein. So these are the given amino acids but, in addition to that phenyl alanine is also there. So they provide hydrophobicity to the protein molecule.

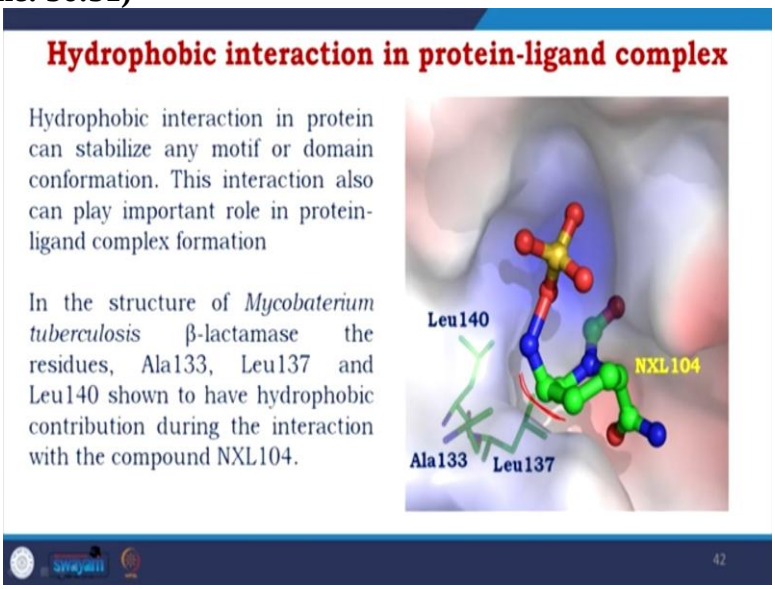
(Refer Slide Time: 56:06)



Why do hydrophobic interaction occurs? There has some argue that the hydrophobic interaction is mostly an entropic effect originating from the disruption of highly dynamic hydrogen bonds between molecules of liquid water by the non polar solute. If you see the leaf of Lotus where if you see that; if you put water it will create a droplet because of

hydrophobic interaction. A hydrocarbon chain or a similar non polar region is patch of a large molecule is incapable of forming hydrogen bond with water. These molecules are also known as low water soluble molecule or hydrophobes.

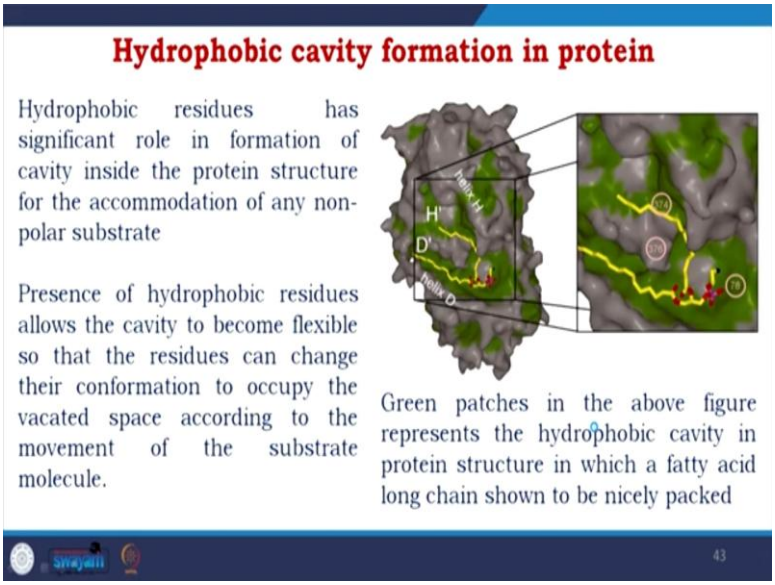
(Refer Slide Time: 56:51)



Hydrophobic interactions play very important role when protein or enzyme interact with a substrate. So, hydrophobic interaction in protein can stabilize any multiport domain conformation. This interaction also can play an important role in protein ligand complex formation. If you look at here, the leucine, alanine, leucine they are interacting with the hydrophobic part of the drug molecule.

So in the structure of mycobacterium tuberculosis beta lactamase, the residues alanine 133, leucine 137 and leucine 140 to have hydrophobic contribution during the interaction with a compound NXL104, the green or the carbon part of the, hydrophobic part of the compound in NXL104 which is also commonly known as Avibactam, a very interesting new generation drug which could have inhibit different beta lactamase enzymes.

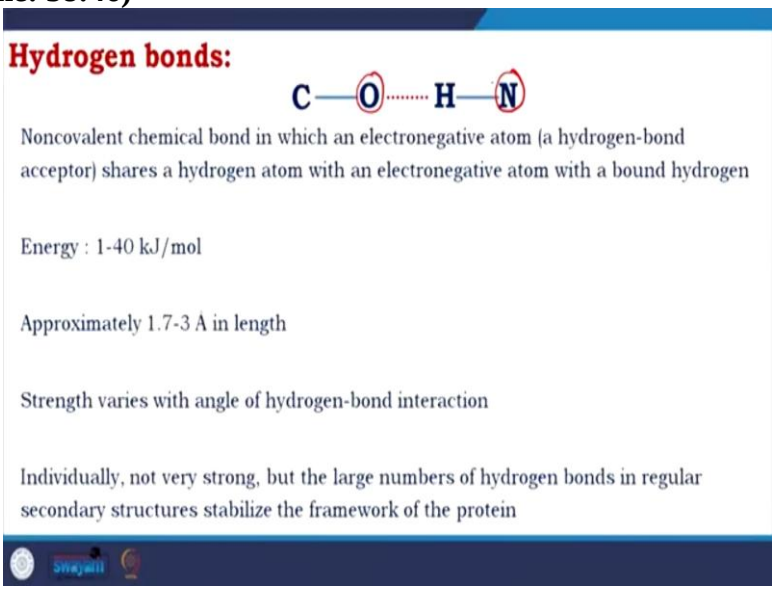
(Refer Slide Time: 57:58)



Hydrophobic cavity formation in protein. So hydrophobic residues a significant role in the formation of cavity inside the protein structure for the accommodation of any non polar substrates. Presence of hydrophobic residues allows the cavity to become flexible, so that the residue can change their conformation to occupy the vacated space according to the movement of the substrate molecule.

So here you could see the green patches in the figure which are hydrophobic cavity in protein structure in which a fatty acid, long chains could be nicely packed. So that is another example of protein ligand interaction by hydrophobicity.

(Refer Slide Time: 58:46)



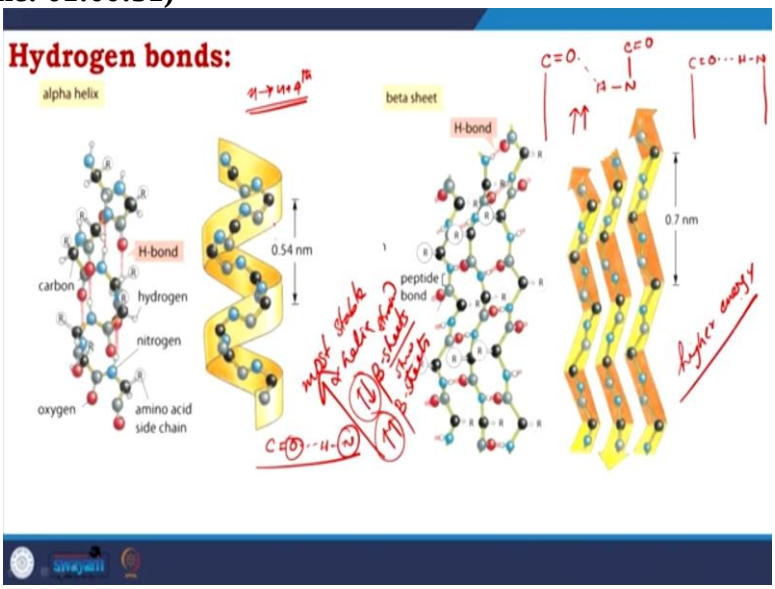
Coming to the next one, the last one hydrogen bond, hydrogen bond, it is a non covalent chemical bond in which an electronegative atom a hydrogen bond acceptor shares a hydrogen

atom with an electronegative atom with bound hydrogen. So mostly oxygen and nitrogen in biology forms hydrogen bonds, energy 1 to 40 kilojoules per mole, approximate length 1.7 to 3 angstrom in length, strength varies with angle of hydrogen bond interaction.

Individually not very strong, but the large number of hydrogen bonds in regular secondary structures stabilize the framework of the protein. So what it is talking about that hydrogen bond probably plays the most critical role in a protein function. As I talked about the protein function needs the change of the protein conformation, so it needs constantly breaking and making of bonds, hydrogen bonds being very less in energy could have form and could have broken very easily.

And by doing that and doing it in a huge number, they help in changing the total conformation of a protein. And that is kind of take a role as a switch on, switch off in protein function which is extremely critical. Also we talked about the strength varies with angle of hydrogen bond formation.

(Refer Slide Time: 01:00:31)



Here, I want to talk about the two very popular motifs, the alpha helix and beta sheets. If you see we already talked about the hydrogen bond between CO and NH is formed. So we talked about the angle dependence of a hydrogen bond, distance dependence we understand, but angle dependence is not understood very easily. Here is a very good example, where we are taking the classic secondary structure motifs which are alpha helix and beta sheets.

So if you see all of them from CONH bond where in alpha helix, it is from n to n plus fourth residue and it is face to face. Here you get the oxygen and the nitrogen in this fashion. Interestingly here, how you get it in parallel? You get it like angled. So this is parallel beta sheets. In anti parallel, you get it face to face, so individually by calculating the hydrogen bond. The most stable alpha helix, second anti parallel beta sheets and third parallel beta sheets.

So stability wise this is the order, but as I talked that beta sheets cannot be staying on the, so I should write it as strands, but as you know how you form a parallel or anti parallel strand, they are already sheet. And that is why they have higher energy. So in terms of energy dismiss leading, individually alpha helix, because of its proper angle and distance from the most stable unit, whereas beta strands are coming together. And it is not possible to measure their energy individually, because individually there is no hydrogen bond.

So we have to measure the entire sheet and that is why they show more energy. So we are done with this discussion. There are so many non covalent bonds. We know about them. But the interesting part is that how they come together, we have studied their individual effects and their physics, chemistry which are according to them. But it would be really interesting how they work together. And that gives rise to a phenomena which is called protein folding. As you see from the history I have talked about protein is studied for more than 100 years.

But and one of the focus of studying protein is to understand its protein folding. If we could understand protein folding, we would be close to immortality. So protein folding is still an enigma and in spite of the efforts of many biochemist, biophysicist, structural biologists. In the next class, we will talk about how protein folding happens, what are the thermodynamic phenomena associated with them and more importantly, how we could study those things? Thank you very much.