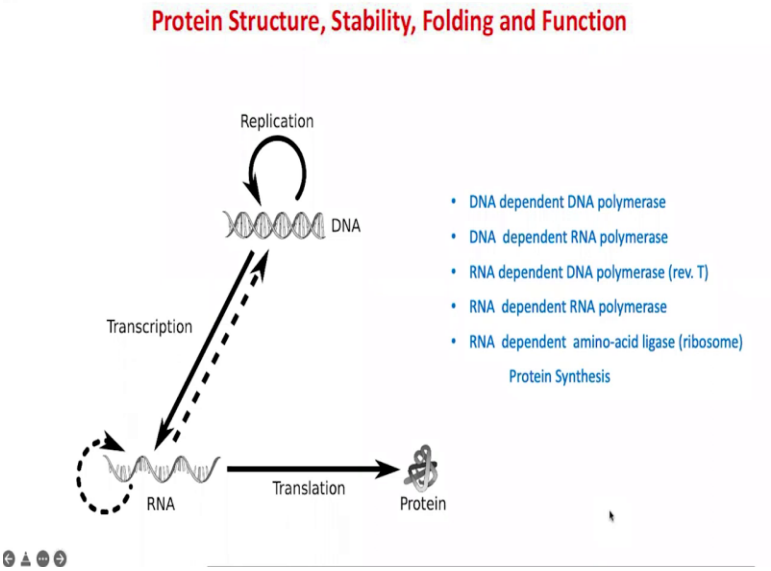
Introduction to Cell Biology
Professor Girish Ratnaparkhi
Professor Nagaraj Balasubramanian
Department of Biology
Indian Institute of Science Education and Research, Pune
Protein Structure, Folding and Function - Part 3

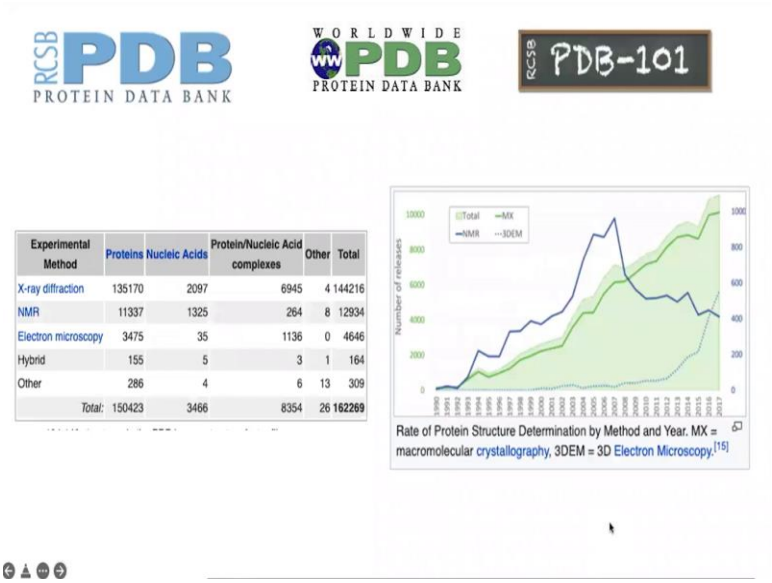
(Refer Slide Time: 00:15)



We are pretty much at a point where you have had a reasonable overview of what atomic level protein structures look like. You also know what atomic level DNA and RNA structures look like. You have seen a lot of animations which gives you a visual idea of what is going on inside the cell. And today, I will be focusing mostly on the, less on the structure, I hope, and more on the stability folding aspects of proteins.

Because as I have repeated again, and again, protein is the last stage of information transfer in the central dogma and proteins make up a huge chunk of our, of the macromolecules inside the cell and they provide a lot of structural and functional roles which keep us alive. Now, before I go into protein folding and stability, a few reminders and some homework.

(Refer Slide Time: 01:15)



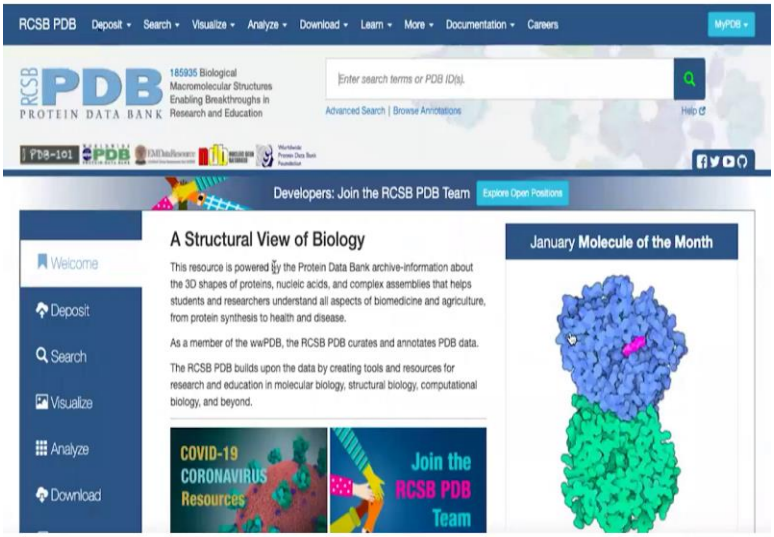
So, I would encourage all of you to go to the website of the protein databank. You just have to type RCSB protein databank in Google and it will take you to the protein databank. In fact, for a minute or so I will take you over there so that you understand how important this is. This is a repository which has existed for almost 30, 40 years, and it collects in all the structures which have been solved through multiple techniques.

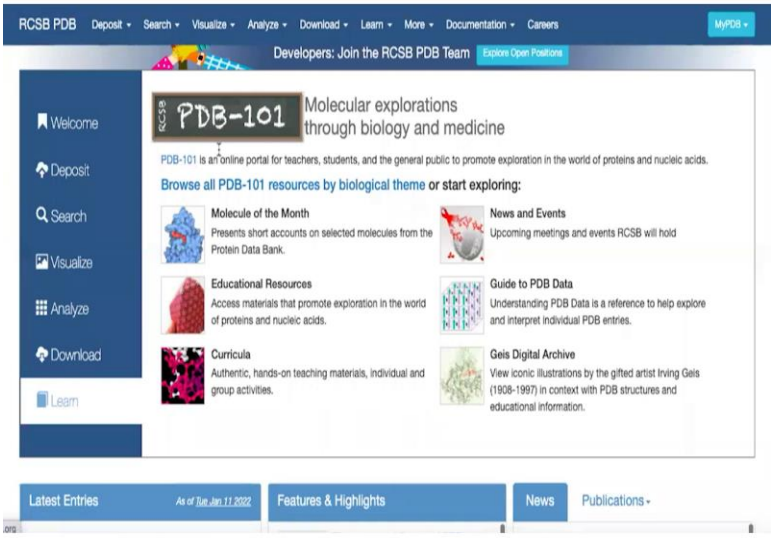
The table on the left shows you x-ray crystallography, nuclear magnetic resonance, the current new hot field of cryoelectron microscopy, hybrid methods, which include modeling methods and others. And it tells you the large number of structures, which have been deposited by researchers all over the world. And these atomic structures are actually maintained in the protein databank and the number is approximately, this is a little old. It is about 1.8 lakh structures as of today.

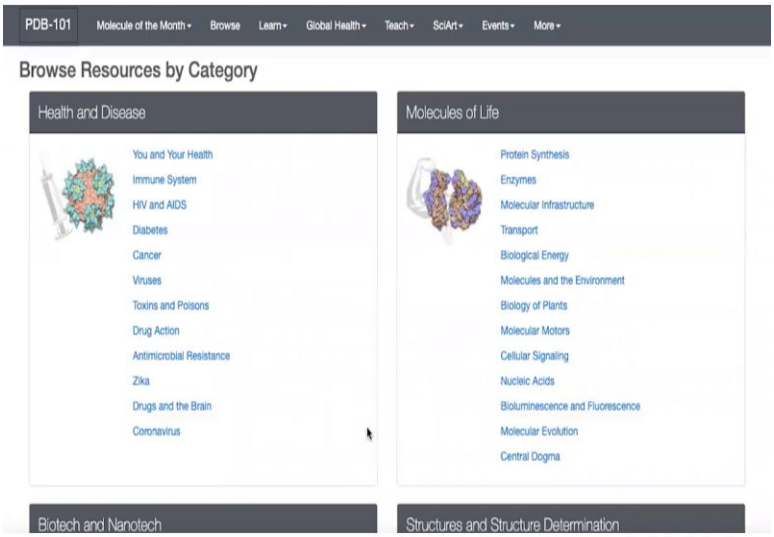
And on the right hand side shows from 1990 when pretty much data and models were being collected till about 2017 the growth of the number of structures which are deposited in the protein databank. And you can see this very interesting idea that nuclear magnetic resonance, which was a very popular technology till about 2005. There is a sudden drop in structures, because NMR has this limitation that it cannot really solve structures of proteins and nucleic acids beyond a certain size. After a certain size, the NMR spectrum which is collected becomes very crowded, and is very difficult to decipher and extrapolate to atomic models.

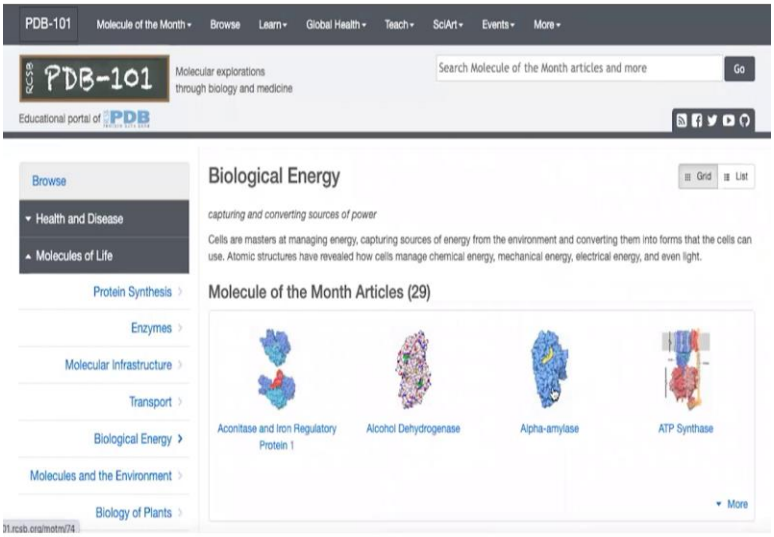
So, NMR as a technique for structure determination, especially of large complex macromolecules is going down. But crystallography, x-ray crystallography continues to be a very popular technique with numbers growing. And you can see this dotted line over here, which is the new structural technique called cryo electron microscopy, and that is literally taking over the world. For those of you who want to do structural biology in the future, there is a good chance that you will never do NMR for structural biology. 80 percent of you will not do standard basic x-ray crystallography and most of you will end up doing cryo electron microscopy.

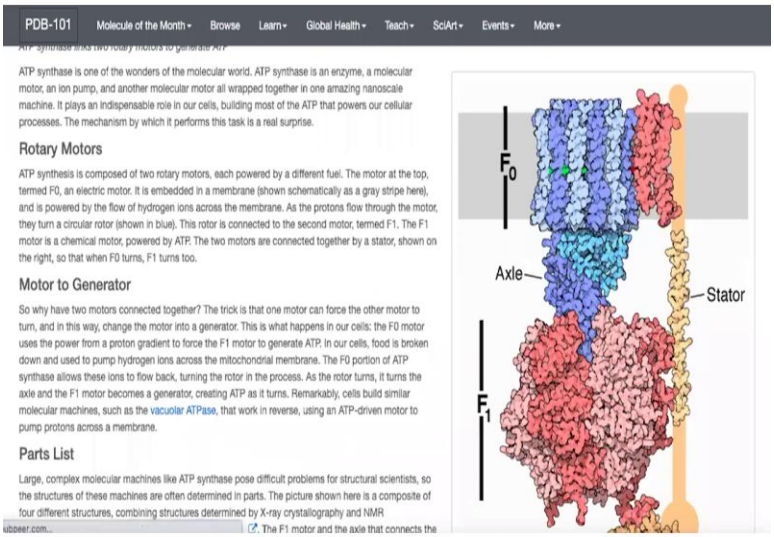
(Refer Slide Time: 03:44)

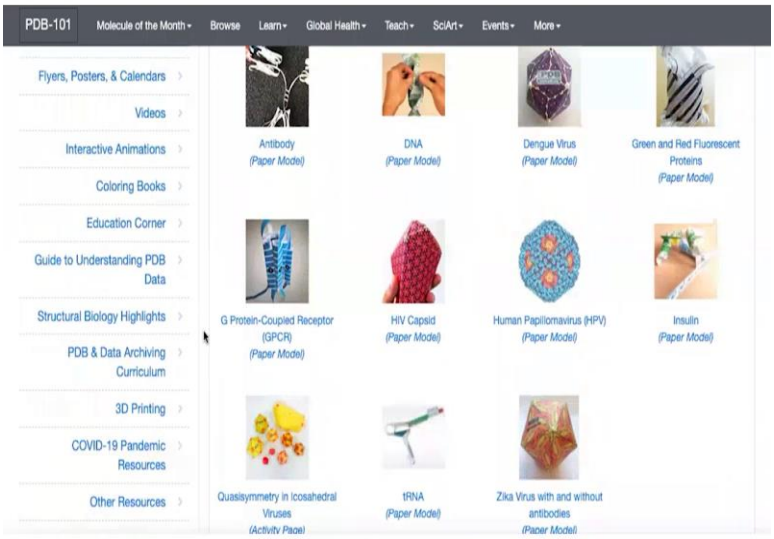





PDB-101
Molecule of the Month
Browse
Learn
Global Health
Teach
SciArt
Events
More
Browse Resources by Category
Health and Disease
You and Your Health
Immune System
HIV and AIDS
Diabetes
Cancer
Viruses
Toxins and Poisons
Drug Action
Antimicrobial Resistance
Zika
Drugs and the Brain
Coronavirus
Molecules of Life
Protein Synthesis
Enzymes
Molecular Infrastructure
Transport
Biological Energy
Molecules and the Environment
Biology of Plants
Molecular Motors
Cellular Signaling
Nucleic Acids
Bioluminescence and Fluorescence
Molecular Evolution
Central Dogma
Biotech and Nanotech
Structures and Structure Determination


PDB-101
Molecule of the Month
Browse
Learn
Global Health
Teach
SciArt
Events
More
RCSB PDB-101
Molecular explorations through biology and medicine
Search Molecule of the Month articles and more
Go
Educational portal of PDB
Browse
Health and Disease
Molecules of Life
Protein Synthesis
Enzymes
Molecular Infrastructure
Transport
Biological Energy
Molecules and the Environment
Biology of Plants
Biological Energy
Grid
List
capturing and converting sources of power
Cells are masters at managing energy, capturing sources of energy from the environment and converting them into forms that the cells can use. Atomic structures have revealed how cells manage chemical energy, mechanical energy, electrical energy, and even light.
Molecule of the Month Articles (29)
Aconitase and Iron Regulatory Protein 1
Alcohol Dehydrogenase
Alpha-amylase
ATP Synthase
More





So, let us just go to the protein databank page. This is the protein databank page rcsb.org. And you can see that it is a very active page. There is a molecule of the month. And with 1.8 lakh molecules and multiple structures being submitted every day into the protein databank, there are plenty of structures to display. There are a lot of exciting structures over here, polymerases, viruses, very interesting shapes and sizes, which I have shown you in the last lecture.

Good place apart from just looking at structures, which is something you can do is to go to the Learn module over here. So, once you go to the Learn module, it takes you to this fairly recent module called protein databank 101. And 101, as you know, is the generic term used for a basic

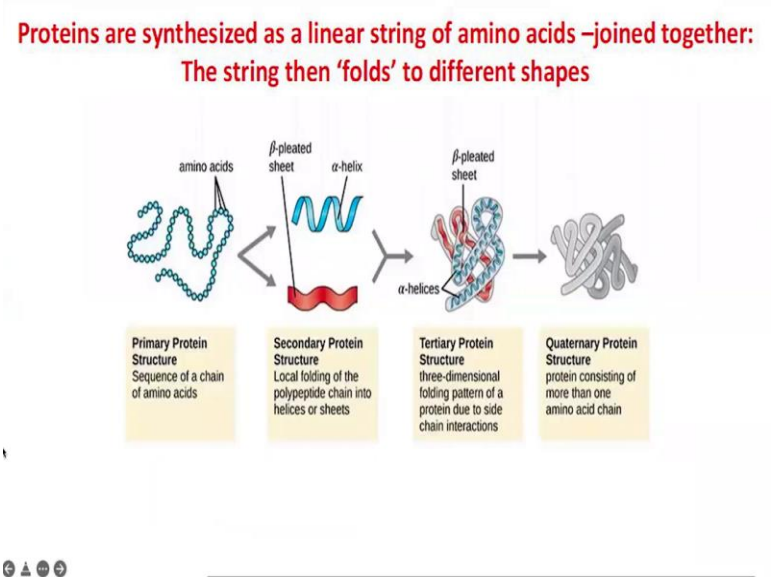
course. For example, the course you are going through right now is a 101 course. And you can use this to go through biological themes you can explore.

For example, let us look at 101 resources by biological theme. So, you have all these themes, which are here, and the Coronavirus will be part of the themes over here. Let us look at something called biological energy for example. And it will take you to all these molecules and remember these are atomic structures. And what you are literally doing is looking at molecules, their models in a representative manner. But these are real molecules which have been solved. These are not models.

For example, this is a famous ATP synthetase, which is sitting on the mitochondrial membrane. And the shape and the size is very, very important to generate ATP. And they are subtleties over here, which you will understand when you do a more higher level course on structural biology so lots and lots of resources over here.

There is also a resource which let me see if I can find immediately, which takes you to a model of a cell. And from that model of a cell, you can basically click different sections of a cell and get an idea of which proteins are there in which section of a cell. So, I would not spend too much time over it. This is basically something you can explore. And again, this is really not part of your curriculum in terms of exams. But for those of you interested, I strongly encourage you to explore the protein databank. So, from here, let us move on to the rest of the course content.

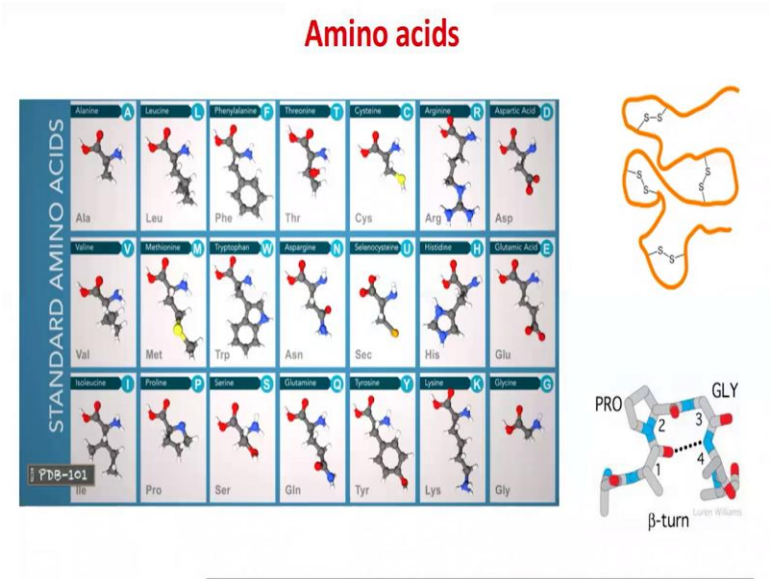
(Refer Slide Time: 06:10)



So, very obviously you now know that proteins as made by the ribosome are a string of amino acids. There are two major secondary structures as defined in protein folding, the alpha helixes and beta sheets. And these two, along with different kinds of helices which are more rare, like a left handed helix, a 310 helix, or different twists of the beta sheets all come together along with turns to make a folded, sometimes compact, sometimes not so compact structure.

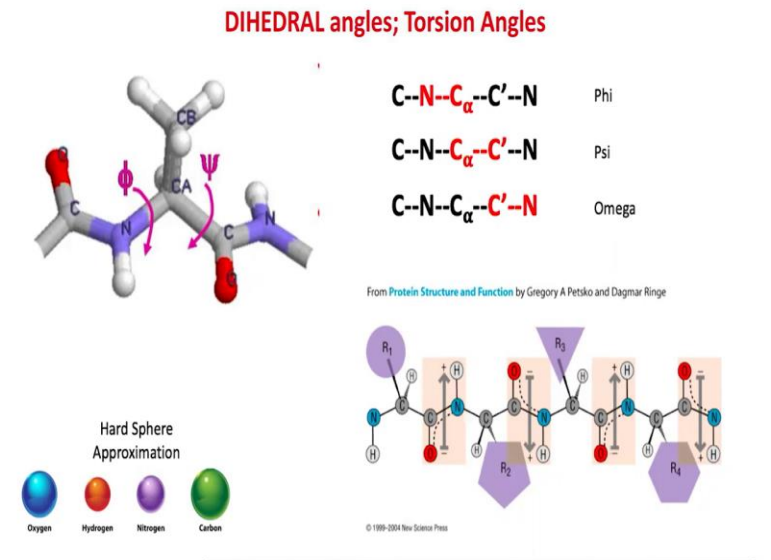
In multiple proteins, sometimes the same protein, like in the case of hemoglobin, sometimes hundreds of proteins can come together to form large quaternary protein complexes. And remember, all of these coming together and have folded states is basically intended to make a functional entity, a tool, which can be used in the cell.

(Refer Slide Time: 07:05)



And you have seen many, many examples of what the shapes and sizes of each functional molecule looks like. Again, this is a ball and stick representation of amino acids in the cell. And here is something I have already talked about. The only connectivity between the linear polypeptide chain is disulfide bonds. And the additional thing which I told you is that there are turns in proteins, and proline and glycine are favorites at the turn. I say favorites, I do not see, I never, I have never said essential. Any amino acid can be in a turn, but proline and glycine do provide some additional advantages because of their properties.

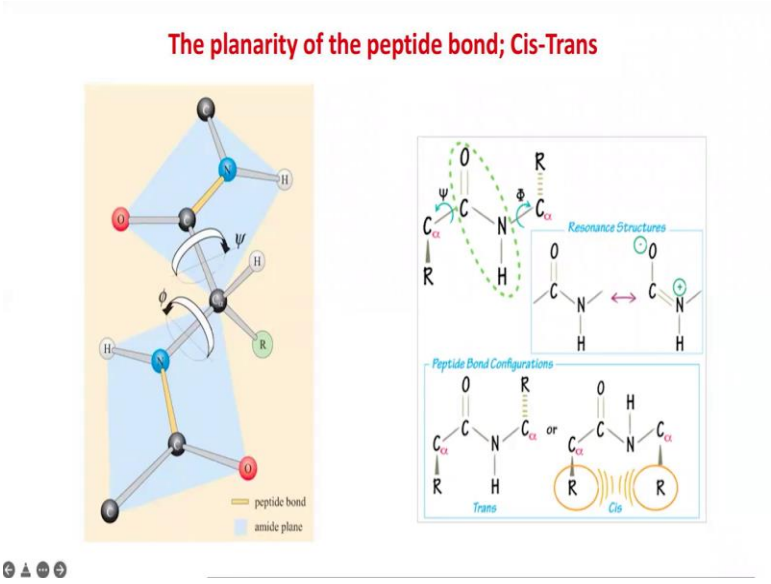
(Refer Slide Time: 07:42)



I have also introduced the idea of dihedral angles, and the concept of phi and psi and told you very clearly that four atoms define a dihedral angle, not three, and phi and psi are have the ability to move. But the omega angle is usually a rigid angle. Because of the partial nature, the electron cloud is dispersed over the CO-NH bond, making a partial double bond over here. So this is a rigid bond, which does not move. It is usually at 180, which does not change beyond a single flat plane, an angle of 180 degrees.

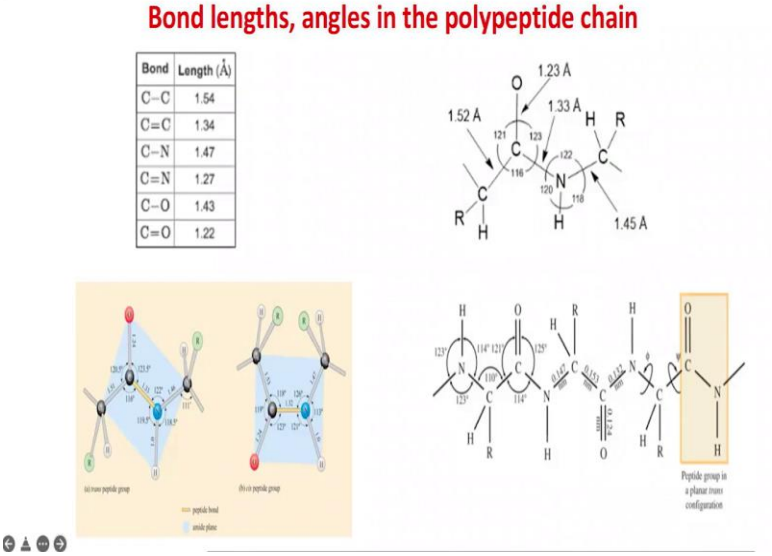
And I told you a lot about hard sphere approximations and how GN Ramachandran did two things. He, of course, made the Ramachandran map. And he also gave a very nice model for collagen based on fiber diffraction pattern. These are the two things he is very famous for.

(Refer Slide Time: 08:35)



Again, a repetition of the fact that the CO-NH bond is basically at a trans configuration, with the R groups facing completely opposite to each other. You have to imagine this in three dimensions. The cis conformation is energetically unfavorable and also based on the hard sphere approximation. The R groups, for example, if you have a tryptophan here and alanine over here they will clash physically with each other. So, cis peptide bonds are almost unknown. The only exception being, and again we would not go into detail, being proline. Proline usually allows the main chain to be in the cis confirmation in about 20 to 30 percent of the cases. It is still not something very predominant.

(Refer Slide Time: 09:24)

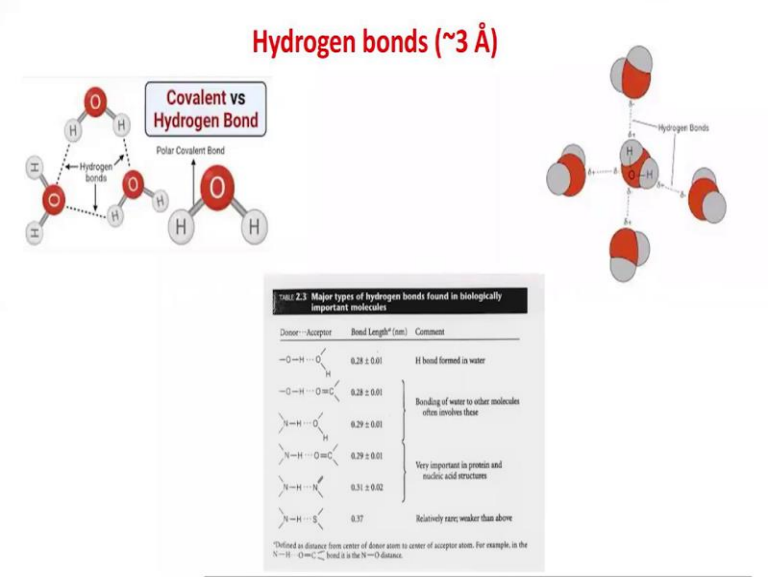


So, today, in addition to looking at the peptide bond, and I have told you this again and again, and this is important for your exams also, please learn to draw nucleic acids, please learn to draw the polypeptide chain, just draw four amino acids one after the other, link to each other and draw the R groups, have an idea about the structure of a few amino acids and practice drawing, because this really helps your brain to understand the different properties of these amino acids.

So, today's focus is on scales, specifically distances. And what is listed over here are the covalent bonds and approximate sizes for these covalent bonds. For example, a carbon-carbon bond is usually around 1.5 angstroms. And carbon-carbon double bond is shorter and stronger 1.3 angstroms takes more energy to break this bond. Here are length units for the carbon-nitrogen, the carbon-oxygen and the carbon-oxygen double bond.

And shown in this picture over here are the distances for these angles. For example, in a peptide, the carbon-carbon is around 1.5 to close to the generic length of 1.54. The carbon-oxygen double bond is the CO here which is part of the peptide group CONH, CO is 1.23 angstroms short 1.3 angstroms over here, and also shown over here are the angles between these atoms. And these angles are of course defined by three atoms, not they are not dihedral angles. So, keep your eye on the basic idea of the length of bonds and let us then move on to water.

(Refer Slide Time: 11:12)



Now, water has very interesting properties, and I will touch upon these near the end of today's class. But more importantly, if you take water in solution, water molecules, independent water molecules are connected to each other via hydrogen bonds. And you saw in DNA structure and you have also got a hint in protein structure, especially secondary structure or tertiary structure, that hydrogen bonding is really very important to provide stability to nucleic acids to proteins.

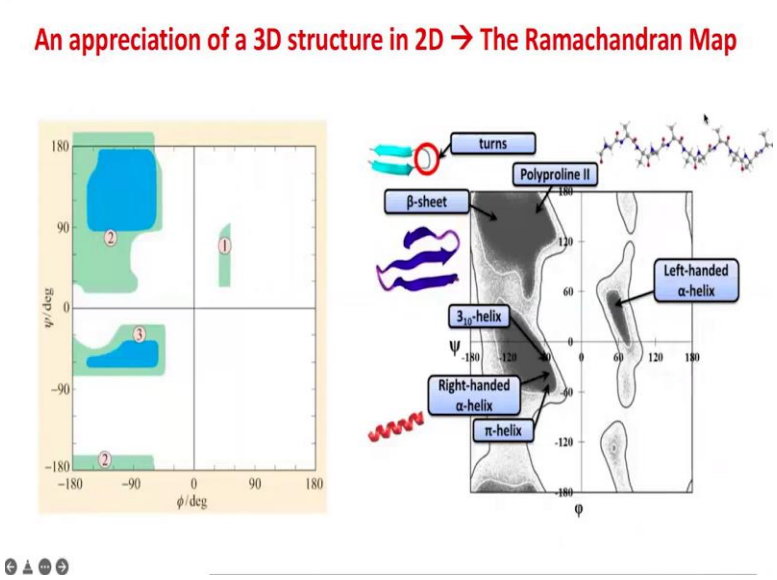
And in pure water, single hydrogen, single water molecule usually is surrounded by, in a dynamic manner by four other water molecules. So, the electrons on the oxygen can form hydrogen bonds with the hydrogen on two opposite water molecules. And of course, the two protons on the water molecule can also independently form transient hydrogen bonds with oxygens in nearby water molecules.

And this hydrogen bonding from the viewpoint of a biologist is very important, because it provides stability to macromolecules, and of course, not necessarily macromolecules all kinds of molecules when molecules are in iron solution. And if you look at the size of this hydrogen bond, it is approximately around 3 angstroms. The lengths here are given in nanometer rather than angstrom so you have to multiply by 10. So, it is a longer bond than a covalent bond. And it is a bond which is fairly easy to break as compared to a covalent bond.

And water molecules, liquid water for example over here, will have dynamic breaking and rejoining of all these water molecules. And when you convert water into ice, this dynamic

structure becomes a little more frozen with the, with these bonds being fairly important in the transition.

(Refer Slide Time: 13:07)

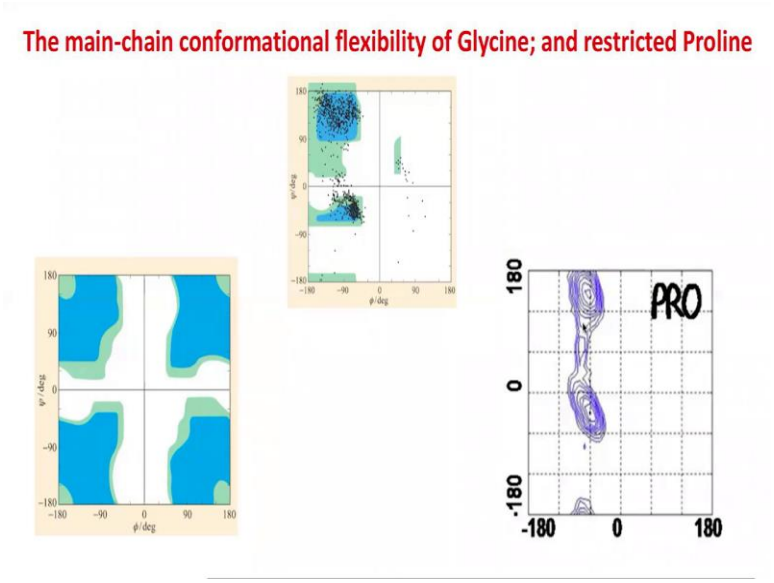


I also told you a fair amount about the Ramachandran map and you now have an idea of torsional rotation about phi and psi. You also know that if you plot a three dimensional protein structure in two dimensions, the three dimensional space which is represented by this quadrant in a graph paper, quadrant 1, 2, 3 and 4, you realize that almost 60 percent to 70 percent of space cannot be occupied based on phi-psi angles simply because there are clashes.

Ramachandran theorize this as hard atom clashes but we look at it in terms of energy these days. And they are only restricted areas in conformational space, where phi-psi angles are possible, and which represent the two dimensional representation of a three dimensional reality which is what is happening in proteins.

And over on the right over here shows you the broad areas where the phi-psi angles are for beta sheets, for 3-10 helices, for a normal right handed alpha helix, for a left handed alpha helix and so on and so forth.

(Refer Slide Time: 14:20)



Now, there are two special amino acids, and in the middle are shown plots for a few 100 amino acids showing you where the phi-psi angles usually are in protein structures in general. You can see there are a few outside the classical Ramachandran map. And part of the explanation for this is on the left, this is a Ravichandran map for glycine. If you take a polyglycine stretch and you ask how much space can glycine occupy in phi-psi space, you realize that glycine unlike any other amino acid can occupy a much more conformational space.

So, you have this symmetric structure with glycine able to occupy conformational space in let us this quadrant on top, on top right and this quadrant on bottom right, which any amino acid including alanine with a side chain cannot, simply because the side chains will clash with each other if you have anything larger than a proton has as a side chain.

Also interesting is proline. You will see compared to other amino acids, all the amino acids, alanine, tryptophan, tyrosine, glutamate, you do not see any dots for a polyproline on the right hand side of the Ramachandran map. So, proline is very strongly restricted because of its side chain being bonded to the main chain. And it is restricted very much. I do not exactly remember what this point is, probably 60 degrees in phi space and slightly broader range in psi space. Sorry, this is phi space and this is psi space, fine.