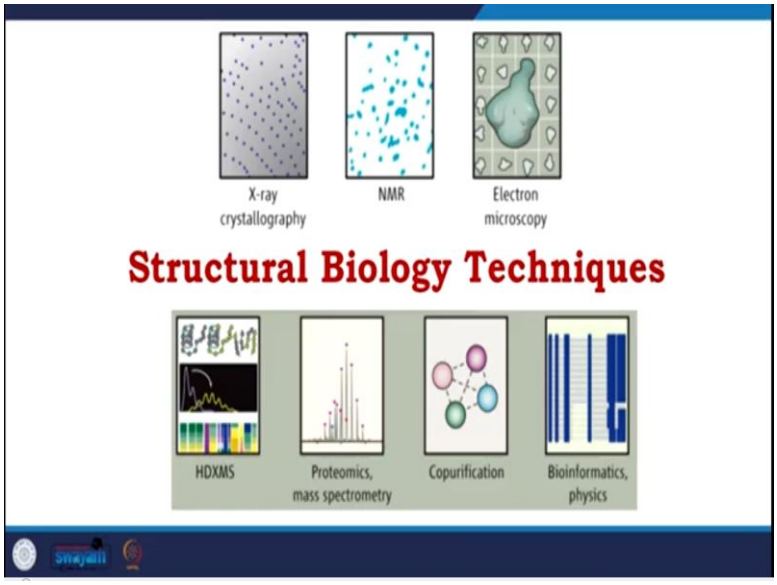
Structural Biology
Prof. Saugata Hazra
Department of Biotechnology
Indian Institute of Technology – Roorkee

Lecture – 11
Introduction to Structural Biology Techniques, Part – I

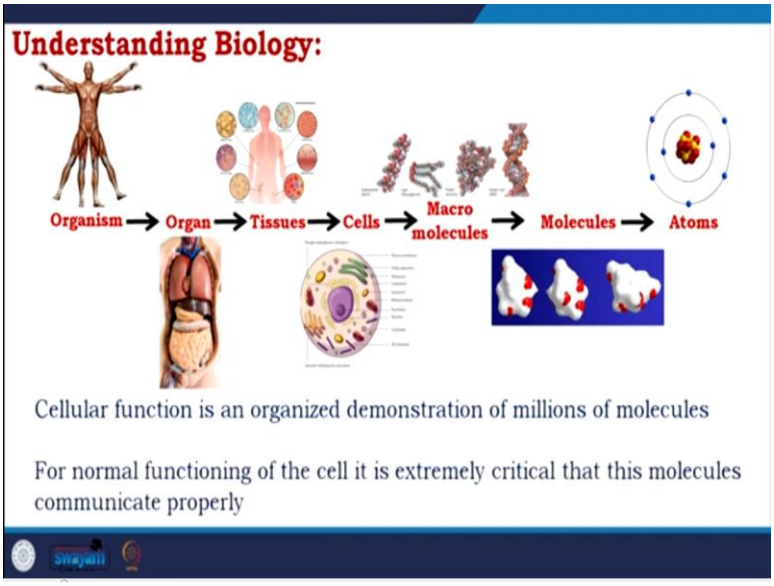
Hi, welcome again to the course of structural biology, we have discussed about backgrounds, why we need to study structural biology, what about the architecture of protein, protein folding and all those things. Now, we are going to talk about the techniques of structural biology.

(Refer Slide Time: 00:55)



In this part as I have already introduced myself to you as a protein crystallographer, we will discuss X-ray crystallography in detail with that I would also make comparative discussion of NMR and Cryo electron microscopy. Also, I will talk about other techniques like CD, like fluorescence, like small angle X-ray crystallography and others, but more importantly, I would also discuss how computational techniques could help solving or analyzing the data in all the techniques. So, welcome to the structural biology technique.

(Refer Slide Time: 01:53)



Before going into the techniques, I just want to talk about the system, we have to understand at the organism level from there, we go to the organ level from there we go to the tissue level from there we are going to the cell level, in the cell there are macromolecules, DNA, RNA, and proteins, and they are composed of small molecule, also there are small molecule metabolites and ultimately, everything comes from atoms and electrons.

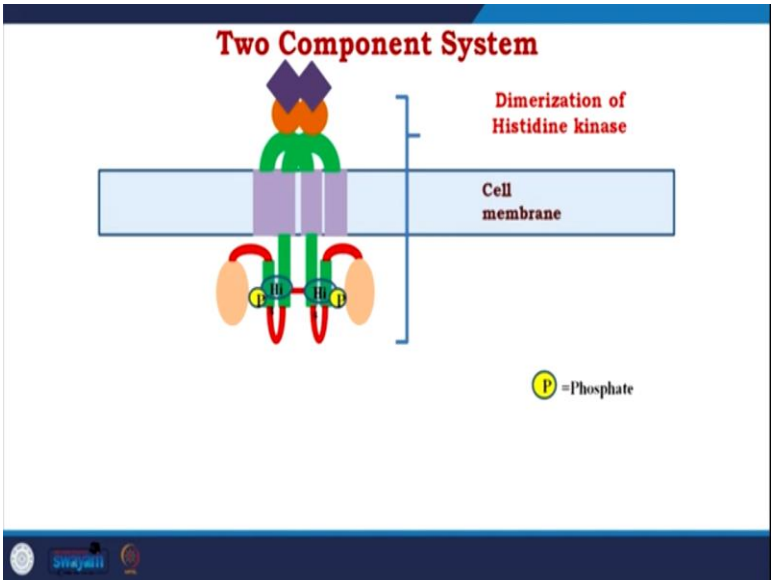
So, to understand biology, we should not be or could not be limited to knowing about the muscles knowing about the bones, we have to go to actually in the atomic level, and it is very important because cellular function is an organized demonstration of millions of molecules. For normal functioning of the cell, it is extremely critical that these molecules communicate properly. And when I am talking about communication, I have chosen to describe you a beautiful communication system in bacteria. And I would also show you how the communication is understood by us using high resolution structural biology techniques.

Let me take an example. Suppose; you are drinking a glass of water and suddenly a drop of water falls in the table where you are sitting. Now, that is a real microscopic situation for you not for the bacteria passing through that table.

(Refer Slide Time: 05:44)

Where the drop actually falls down, for that bacteria suddenly there is a sea of water created, and the bacteria was not even prepared for that. So, that is the problem of microscopic organisms. And because of that, they have a total system which helps them surviving from emergency situations. One of them which, is very important is called two component systems.

(Refer Slide Time: 06:27)

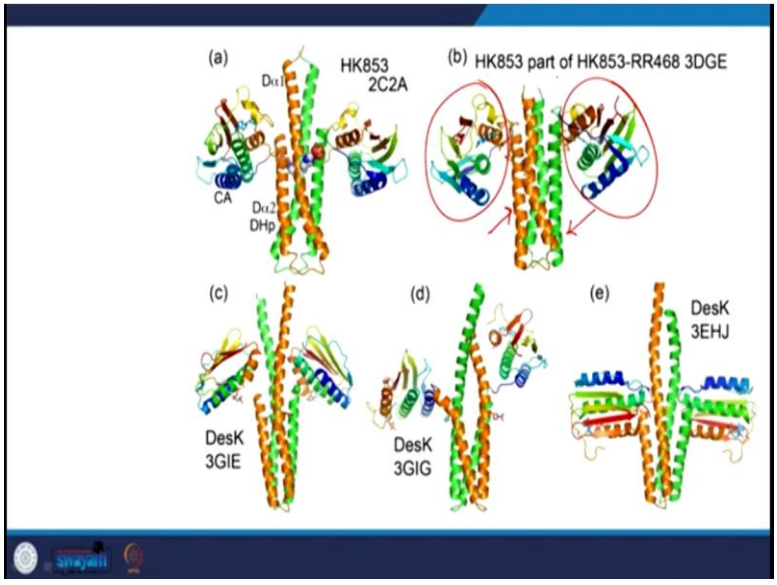


A two component system as the name suggests, is made up of two proteins, one is a Histidine kinase and other is a response regulator. Histidine kinase is a protein which actually stays on or in between the surface of the organism. So, it has a part which exists in the periplasmic domain, it has another part which is in the transmembrane domain and a kinase domain in the cytosolic part.

The periplasmic domain, it works as a receptor, there are different types of receptors, responding to different emergency situation. When this receptor received an external stimulus, it respond to that and convey the news to the bacteria, and help bacteria expressing the proteins which could help them to survive. For example, if the external stimulus is sensing heat, it would communicate to the bacteria and bacteria will express heat shock proteins.

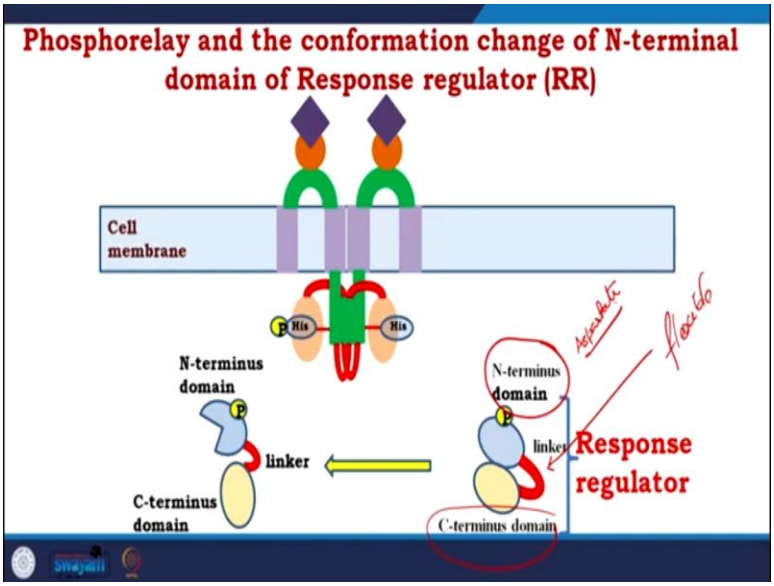
So, how they communicate is beautiful and very interesting, whenever this external stimulus bind to the protein Histidine kinase, this Histidine kinase dimerizes and this dimerization help them to bind a ATP molecule. When the ATP molecule bind, they auto phosphorylate in the Histidine.

(Refer Slide Time: 08:56)



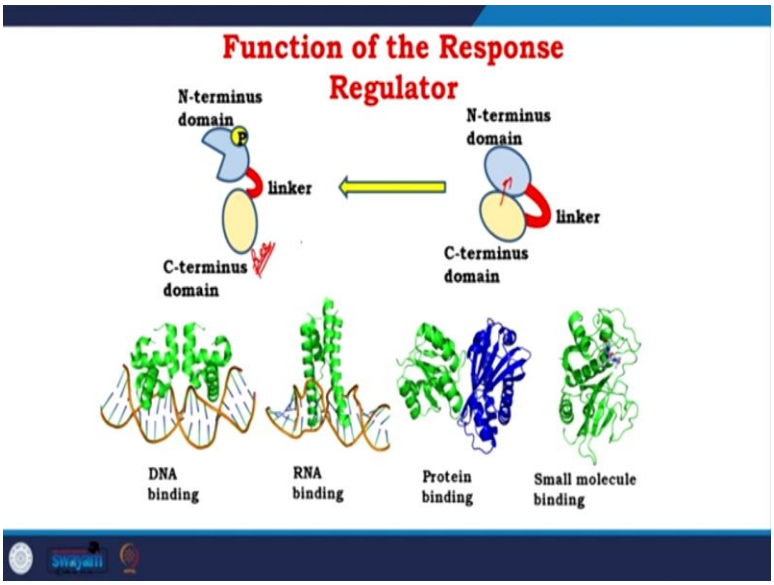
But let us take a look at the high resolution structure of these Histidine kinases. You will see that they have the dimerization domain made by this orange and green color alpha helical domains, these are the kinase domains, you will see mostly similar architecture of protein in all the Histidine kinase.

(Refer Slide Time: 10:08)



So, coming to the next part, now the phosphorylated Histidine transfer the phosphate to response regulator which is having two domains. N-Terminus domain and C-Terminus domain connected by a flexible linker, the phosphate comes and bind to the aspartate residue of the N-Terminal domain and that leads to a changeup conformation and opening up between the N- Terminus and C-Terminus.

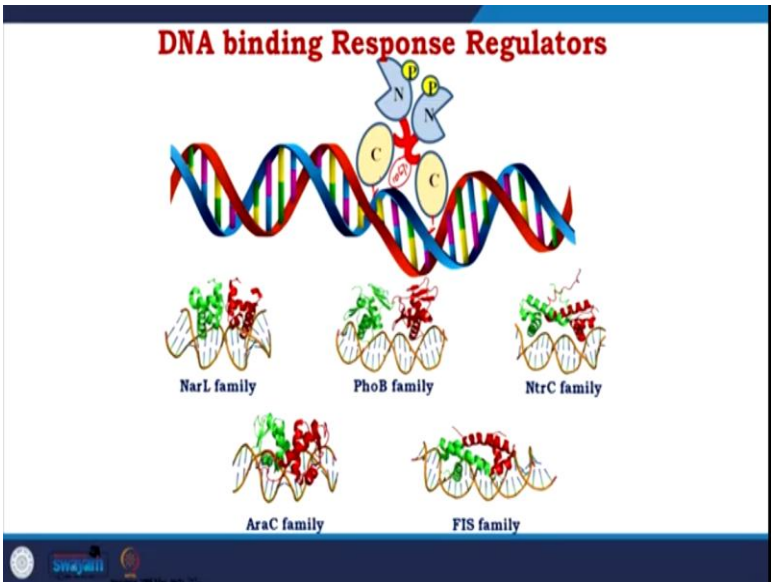
(Refer Slide Time: 10:57)



When this change of conformation happens, the C-Terminus which was initially bound to N terminus is now free. And being free it interact with many macromolecules and small molecules like to a DNA to RNA, to protein, to small molecule. So, by binding to different biologically

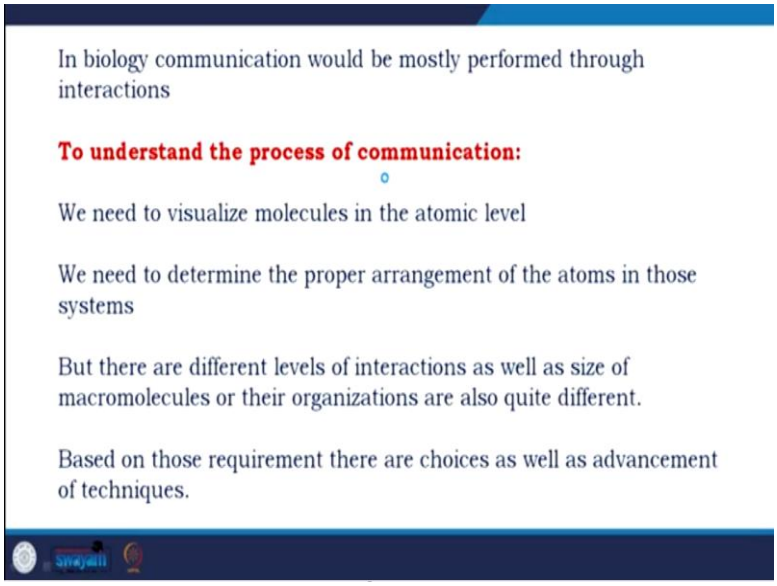
relevant molecule, it actually relay the signal, the Histidine kinases have received through its receptors.

(Refer Slide Time: 11:47)



If you look at the DNA binding, you would see that, the DNA binding is very symmetrical. There are different families, if you see NarL family, PhoB family, NtrC family, AraC family, FIS family, they are all DNA binding protein, but if you closely take a look, you will see that the binding patterns are very different. Also, you would be amazed to know that they could also control the type of signal. So, depending on the binding affinity, depending on the way they are binding, they are actually relaying different signals. Just for a simple example, I talked about heat, it could be low heat, it could be like medium amount of heat it could have a very high amount of heat. Depending on them they choose target, they change binding affinity and in that way they do the relaying of the signal. And all these stories I am talking about could be understood in terms of solution of high resolution structure. So it is essential in biology to look at atomic level, if you want to look at the intricate mechanism of communication they are making.

(Refer Slide Time: 14:00)

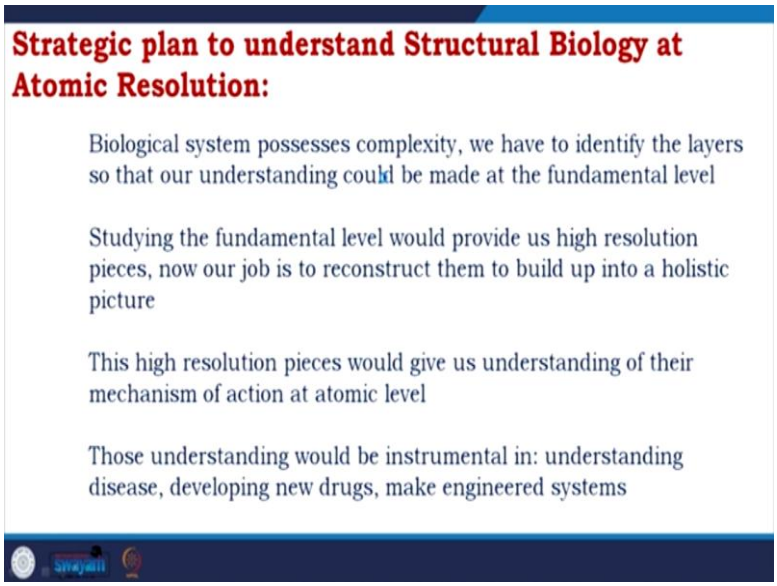


So, in biology, communication would be mostly performed through interaction. To understand the process of communication, we need to visualize molecule in the atomic level. We need to determine the proper arrangement of the atoms in those systems. But there are different levels of interaction as well as size of macromolecules or their organizations are also quite different.

What I mean is, if you think you want to look at one Histidine kinase, and as I have shown you, you got it in the cytosolic portion. have certain size, depending on the molecule. They are like 300 to 500 amino acid let us say, but now, if you want to get the dimer, then it would be 1000 amino acid. If somehow you want to get the entire thing with the membrane, it would be a much larger assembly. Again, if you want to get together the Histidine kinase along with the response regulator, they will be even bigger, if you want to get in the next level, the Histidine kinase, the response regulator and the molecule it is binding, it even create bigger assemblies, not to stop with that, sometime I talked just about dimer, they form Multimars. So, in biology, most of the structural techniques we will be talking about atomic level, but at the same time, larger assemblies like an entire virus, today, we are in the situation where COVID is a big problem.

So, people are trying to get the structure, people are solving the structure of entire virus. And with that level of organization with that level of different side's difference, we need different techniques. So, based on those requirements, there are choices as well as advancement of techniques. I will talk about Cryo electron microscopy, electron microscopy when started, was actually in much lower resolution. But today if you look at, this is sensation, this is getting into atomic resolution, we will talk about that later.

(Refer Slide Time: 17:05)



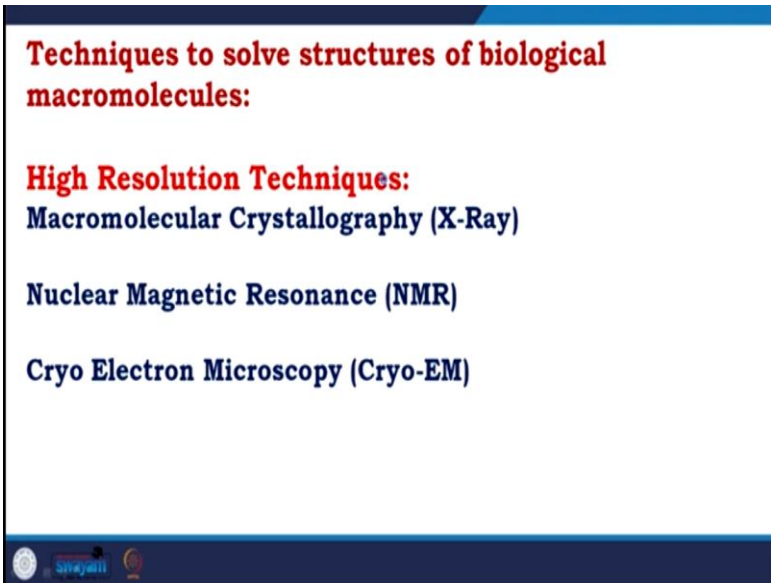
So, what should be the strategic plan to understand structural biology at atomic resolution, biological system poses complexity; we have to identify the layers, so that our understanding could be made at the fundamental level. Initially, the challenge was, we want to understand the structures of the molecule, the simple molecules like a single protein. Now, with the advancement of technique, with the dedication of great scientists, we have achieved a lot. Every day solve new structures, but understanding of biology is not being limited to that. Now, we want more, we want from one protein to multiple assemblies, from multiple assemblies to the molecular masons and so on. Studying the fundamental level would provide us high resolution pieces; our job is to reconstruct them to build up the holistic picture, so that we could understand the function at the level of a single protein.

But at the level of machinery like how a DNA polymerase is working with the whole replication complex or transcription complex or what we got is the ribosome in the translational complex. This high resolution pieces would give us an understanding of their mechanism of action at atomic level, those understanding would be instrumental in understanding diseases, if we could understand how it is happening.

We could also understand, how the abnormalities coming? Let us say in a cancer disease, we understand which protein loses control, which protein is expressing without control, and how that is making abnormality? At the same time, we will know the structure of the protein of viruses or bacteria, which we want to target as a drug development target also, now, this is the

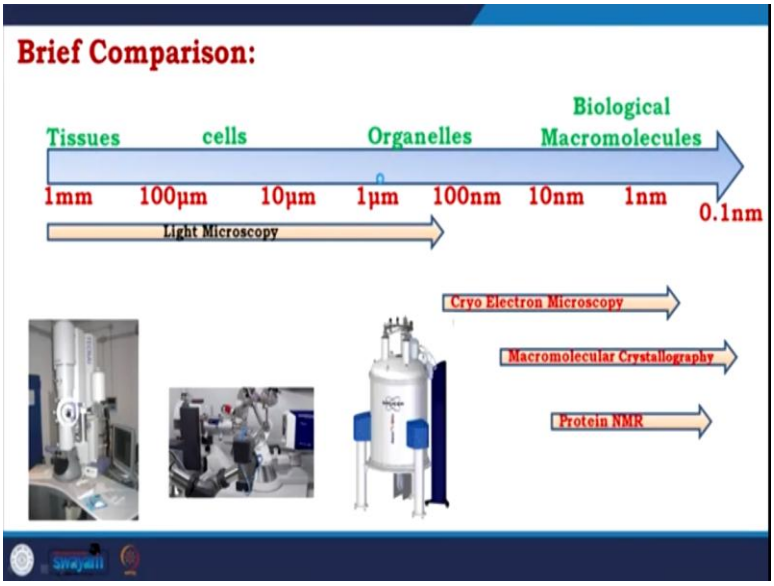
arena of engineering, a lot of protein engineering, metabolic engineering, enzyme engineering is happening. Very recently, Francis Arnold was got Nobel Prize for protein engineering. So to do engineering, again we need to look at the atomic level that is the necessary condition.

(Refer Slide Time: 20:04)



So, techniques to solve structures biological macromolecules: High resolution techniques, macromolecular crystallography X-ray, as I told I will talk about nuclear magnetic resonance or NMR, Cryo electron microscopy or Cryo-EM.

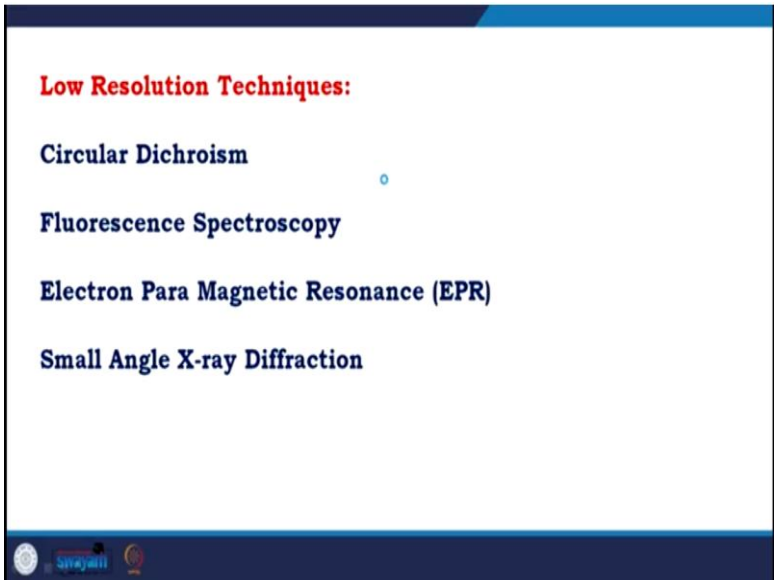
(Refer Slide Time: 20:27)



So, let us take a brief comparison of them. If as I started with, we look at the size starting from tissues to our atoms, it ranges from 1 millimeter to 0.1 nanometer. Light microscopy, which is a

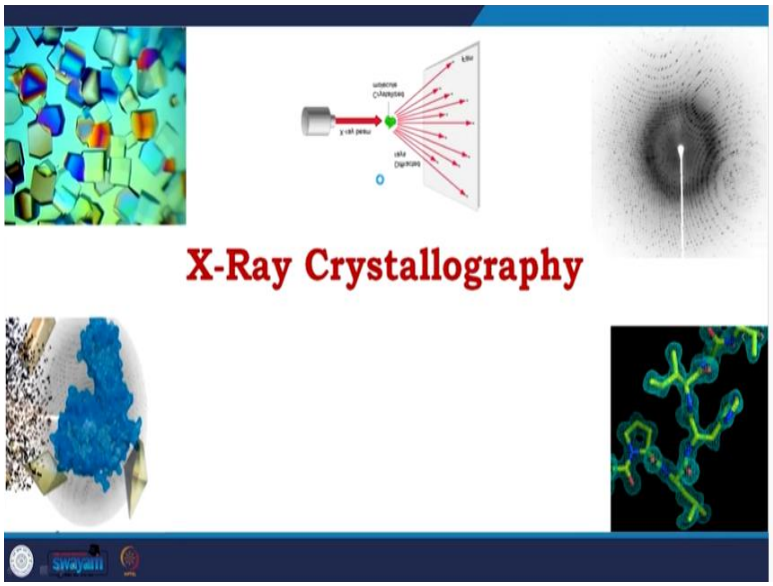
real fundamental tool for us to look at lots of things, helps us to look at from the level of tissues to the level of organelles, but beyond that, it is not possible for light microscopy. We have Cryo-EM, Cryo electron microscopy, which starts around 100 nanometers and currently goes to the angstrom level. Cryo-EM is a new generation sensation and I will talk about that in more detail when I will talk about Cryo-EM especially. X-ray crystallography as I told I want to talk about that in detailed with the entire process. X-ray crystallography or macromolecular crystallography contribute mostly to the structure solution up to now, nearly 90% of the solid structures are done by X-ray crystallography, it ranges from around 10 nanometers to angstrom level, good structures are considered from 3.5 to 1 angstrom. NMR has relatively lower range, it cannot work with bigger molecule because it works with the tumbling, but NMR has its own applications especially when it considers dynamics, the dynamic property of macromolecule. NMR is really contributing and beside protein NMR have a universal usability which also we will talk in details.

(Refer Slide Time: 22:43)

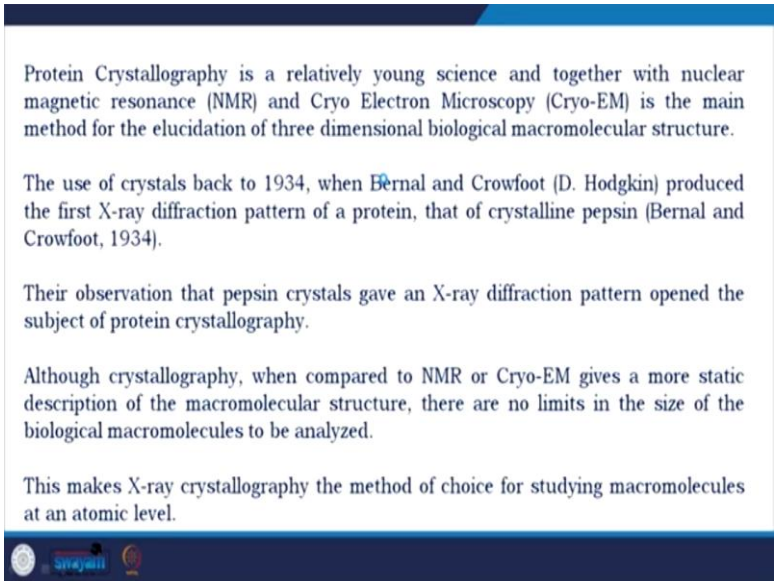


Low resolution technique: There are many techniques I will talk about Circular dichroism, Fluorescence spectroscopy, Electron paramagnetic resonance and Small angle X-ray diffraction, but when I say low resolution, it seems these instruments or these studies are less efficient than high resolution. But these techniques have another set of usability application towards getting solution level interactions, we will talk again.

(Refer Slide Time: 23:31)



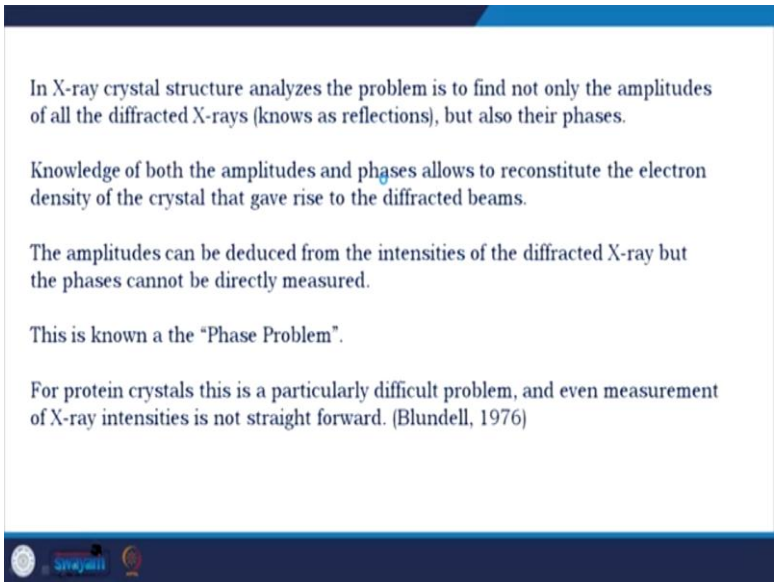
(Refer Slide Time: 23:38)



Today I will start X-ray crystallography. We will talk in details protein crystallography. It is a relatively young science and together with nuclear magnetic resonance, NMR and Cryo electron microscopy or Cryo-EM. It is the main method for the elucidation of three dimensional biological macromolecular structures like protein, like DNA, like DNA protein complex like virus. The use of crystals back to 1934, when Bernal and Crowfoot and Dorothy Hodgkin produce the first X- ray diffraction pattern of a protein, that of a crystalline pepsin and Bernal and crowfoot were awarded a Nobel Prize. Their observation that pepsin crystal gave an X-ray diffraction pattern open the subject of protein crystallography, before people get to look that the

proteins when it is in high concentration, some protein naturally make crystals. Crystals are good. Crystals are easy to purify the protein. So they were used in that manner. But now it says that Crystal also could have provided some meaningful information from its architecture although crystallography when compared to NMR, or Cryo-EM, gives a more static description of the macromolecular structure. That is why crystallography is always criticized. There are no limits in the size of biological macromolecules to be analyzed, once they are crystallized, I will talk about this in details. In crystallography, the main barrier is crystallization, a protein which is in solution would make a crystal or not. This makes X-ray crystallography the method of choice for studying macromolecules at an atomic level, you really get high resolution structure, and you really get to see the molecules in real atomistic vision.

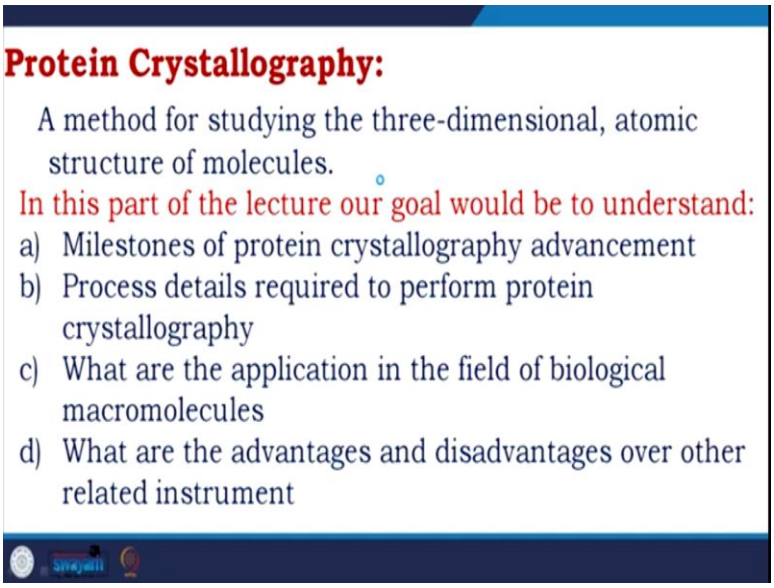
(Refer Slide Time: 26:04)



In X-ray crystal structure analyzes the problems is to find out not only amplitude of all the diffracted X-rays (known as reflection), but also their phases. So, we need these two things, we need the amplitude as well as phase. Knowledge of both amplitude and phase allows reconstituting the electron density of the crystal that give rise to the diffracted beam which will convert it to model. The amplitudes can be deduced from the intensities of the diffracted X-ray, but the phases cannot be directly measured. So, you have a crystal, it is an array of molecules, you hit it with an X-ray, and when the X-ray hit the atoms, then you get intensity and you calculate amplitude. But if you think any Ray or diffracted beam with wave properties, it could be destructive Phase or constructive phase. So, calculation of phase is very important, but it

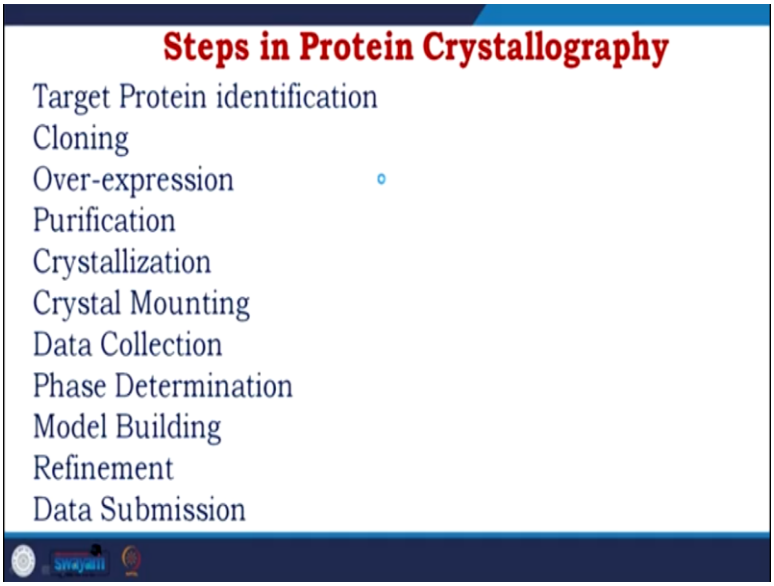
cannot be done through the normal X-ray crystallography experiment. And that is why it is called the famous Phase Problem. So, for protein crystal, this is a particularly difficult problem and even measurement of X-ray intensities is not straightforward.

(Refer Slide Time: 27:55)



So, in summary, protein crystallography is the method for studying the three dimensional atomic structures of molecules. In this part of the lecture, our goal is to look at the milestone of protein crystallography advancement. It is a young science, but it has a lot of achievement around 30 plus Nobel laureates, dedicated their life in doing research using X-ray crystallography or protein crystallography as a tool. So, new generation listeners if you have higher ambition, this could be your technique. We will talk about process details required to perform protein crystallography, from choosing a particular gene target to get a pure protein; this is equally applicable for NMR, Cryo or even with the low resolution structure solution techniques. We also talk about the application in the field of biological macromolecules. And we will compare with NMR and Cryo electron microscopy (detailed comparison to understand what are the advantages and disadvantages of these three techniques).

(Refer Slide Time: 29:54)



Coming to the steps in protein crystallography: Your first work is to choose a proper protein target, but go back to the gene sequence, look at the gene sequence, and then do some engineering. Next is cloning, then overexpression, then purification, for structure solution techniques purification is extremely critical, because you could only get good information if you have very pure protein in solution.

Then, coming to the topic, coming to the part I am talking about from the start to the end, which is crystallization, the biggest barrier in the success of protein crystallography. Mounting the crystal, how you take the crystal from the drop, where we are going to set it to the diffractometer, how you collect data, what is your strategies, how to optimize data collection strategies, then the phase problem I talked about. How to determine the phase? What are the techniques, what the problem challenges are, and how to solve phase problem? Once you get the amplitude and phase, you will go for model building, refinement through energy minimization, and all and once you have the refined model, you will check the quality and ultimately, you will do the data submission. Data Submission as I told in protein data bank is essential. Because once you submit the data (you have to submit the model as well as the electron density), you will receive a validation certificate from them. And once you got a validation certificate from the PDB, then only you are supposed to communicate the journal which is the target of the entire work to publish it so that the international community of scientists could have access to this structure.

With this, I will finish today. If you have any queries please write to us more responses will coming from you. It would be helpful for us. And as I started with, I continue saying this. One of my inspirations to continue this course is to make a good generation of structural biologist from our next generation. Structural biology is becoming more and more critical towards biomedical engineering, drug designing, pharmaceutical, technological problem and whatnot. So, more people would be getting interested, more pupils who would be getting educated. We will be happy with that. We want that. Thank you.