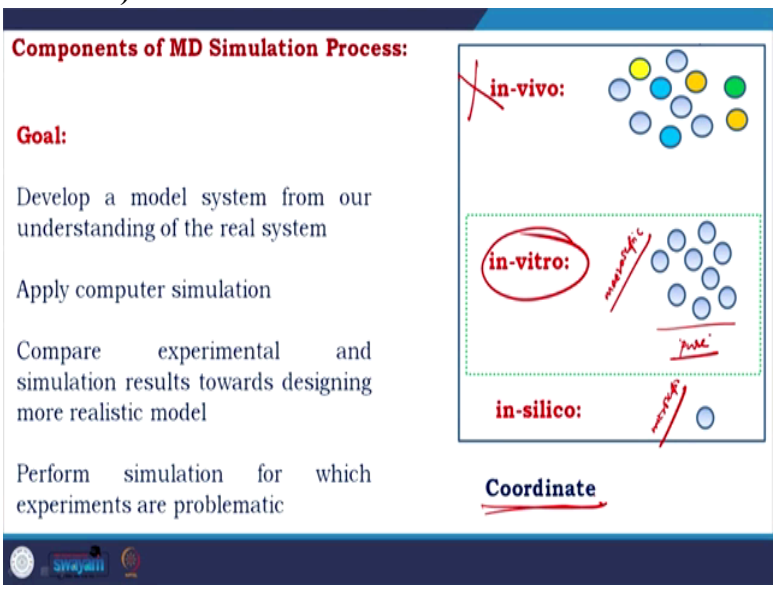
Structural Biology
Prof. Saugata Hazra
Department of Biotechnology
Indian Institute of Technology, Roorkee

Lecture-47
Molecular Dynamic Simulation Process Part I

Today, we are going to discuss the molecular dynamics simulation process. I have divided the whole process description part into 2 classes. So, in today's part one, when we started discussing molecular dynamics, I told molecular dynamics is a process that is held by a computer simulation, and what is happening is we are looking at the phenomena and the parameters which are being taken care of by a real experiment. So, we have a real system and develop a model system. Then the model system undergoes simulation by taking care of their microscopic properties.

(Refer Slide Time: 01:32)



Now, what is our goal here? We aim to develop a model system from our understanding of the real system. As the first goal, you have an experimental system, develop a model system, and then apply computer simulation. For a long time, people are dreaming of developing systems model systems out of experimental systems, but the dream started coming through with the advent of the computer because now you have a huge ability to process data.

You will assume a lot of hypotheses and do the calculation according to that. So, that is where the computer simulation is coming from, and then, you perform an experiment, you perform a simulation, compare these 2 data, and develop a more realistic model. Though this is a never-ending process, a better model system will develop the more you explore. But once we

get good results from the simulation, we use this simulation to perform problematic experiments.

I have explained what I mean by problematic in your common sense, and you can understand that if you have to experiment with real life, it will take time and money, it will take manpower, and it will decrease many things. But if you could have a good model realistic model system, you could perform simulation, and I talked about that when you compare the experimental systems in-vivo, which is inside the body if we consider what we are talking about a protein.

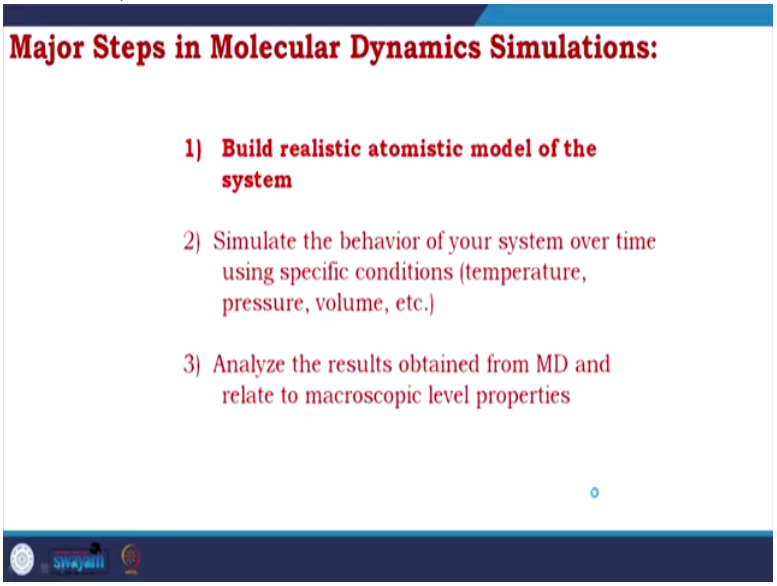
You will see that in in-vivo, there is a mixture of proteins in-vitro. If you have done correctly, the protein purification, you still get your protein in a million copies. Whereas in in-silico, which is the model system, you could have chosen the number of protein molecules according to your choice. That is the best thing for this experiment. We will explain in many cases why that is advantageous. But to tell you a very important thing.

As I told when we are developing a model system, we must mimic an experimental system. In the case of molecular dynamic simulation, we do not mimic in-vivo. Unfortunately, our technological improvement did not reach the level where we could deal with so many components. So, what do we do? We mimic the in-vitro condition. So, in in-vitro conditions, even if not 100%, the protein could be pure, more than 95%.

So, you have a major portion of your protein behaving in a similar identical way. Now, you experiment using what we say pure protein. So, if you look at these, we are called pure protein, and the data would be compared with the simulation result, whereas I repeatedly talk about the number, this is the macroscopic system, and this is the microscopic micro-scale, and then you have the coordinate.

So, when you are performing a simulation, you control the mass and the position, and then what you do, you apply the force field. In the previous class, we talked about different force field components. So, you develop your force field equation by having these bonded and non-bonded interactions, the 6 parameters we have discussed. So, the force field equation will help you to develop the force field. You have the protein molecule, apply force, and see the changes.

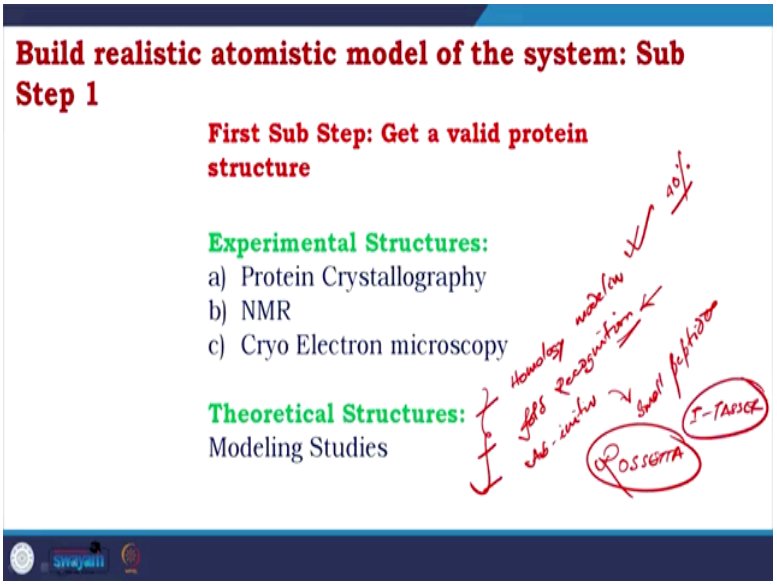
(Refer Slide Time: 06:57)



So, coming to the process now, the major steps where I finished yesterday, there are 3 major steps 1 build a realistic atomistic model of the system. 2 simulate the behavior of your system over time using specific conditions, temperature, pressure-volume, etc. and 3 analyze the results obtained from MD and relate to macroscopic level properties. So, first, you make the initial sample. If you compare it with the experiment as we have discussed in your experiment unit sample preparation, here also we are preparing the sample.

You would be slightly surprised because I was already told that getting the system means getting it from the experimental structures in the last class. So, what else would I explain today? What more do you have to do once you get the coordinates of the PDB file?

(Refer Slide Time: 08:12)



So, the first step we already discussed is getting a valid protein structure. That is all that we are studying in the structural biology course, how to get our journey from the sequence up which we get in huge amounts, millions, and trillions with the advent of NGS, our journey from this sequence to structure now, we have structured initially we told that once we get structure, we could understand the function, but when we start getting structure, we see that not always structure is giving us the function.

So, we want to introduce dynamics; as I told pictures could be incomplete and wrongly projected. So, we are going for dynamic. So, we need to get a valid protein structure, which is the outcome of a structural biologic technique. We know the experimental structures could be obtained using protein crystallography, NMR, and cryo-electron microscopy.

At the same time, we also know that the number of structures we have obtained, which is deposited in PDB, is 1,70,000 only, and even it is not representing 1,70,000 unique proteins because there are important proteins which are having their structure in multiple forms with substrates, inhibitors in different organisms and all these things. So, around 30 to 40,000 unique proteins we have.

What else we always talk about this bridge here we are including the theoretical structures, which we have solved using modeling studies, this course is not the best place, but as it is connected, I keep saying about there are majorly 3 techniques 1 homology modeling, 2 fold recognition and the last one is Ab initio. Now, Ab initio is mostly applied for small peptides, but fold recognition is still away from success, as it is not a proven method for the situations when needed.

Homology modeling is an established technique, especially when the percentage of identity is more than 40%. But, very interestingly, now, there is a combination of techniques, and why I am talking about this combination of techniques at this juncture of the course because, if you look at and I will probably introduce a small part as a case study, there are 2 programs called Rosetta and I-Tassar their program packages and if you see or if you give a detailed look to that.

You will find that the package is the proper culmination of domain modeling, Ab initio, fold recognition, energy minimization, and statistical analysis. So, when we study MD simulation,

you will see how molecular dynamics simulation is also critical in the theoretical solution of the structures. You all know alpha fold certainly got a very interesting success, and we are all looking forward if you look. One of the biggest help they got was through simulation.

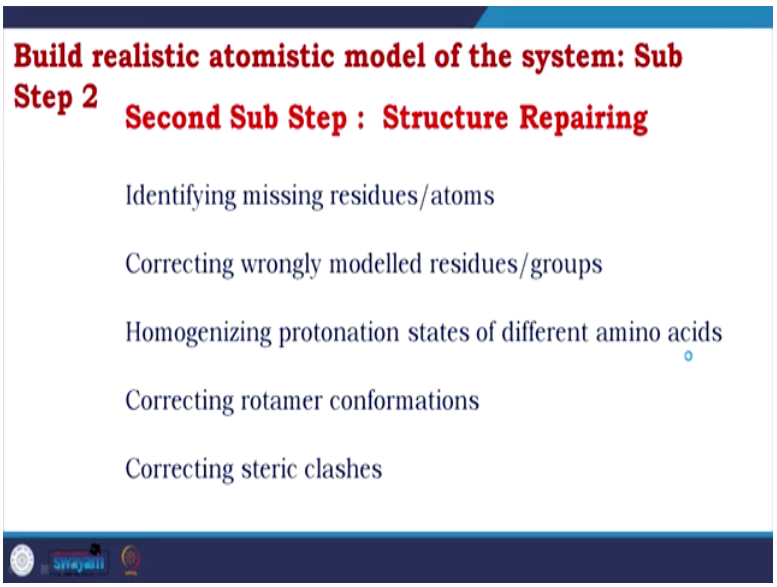
We will talk about the second step is structure repair. Now, this is where you could take a lot of help from people like me, who have solved structures directly by their hand and also work on theory. From the crystallography theory, the first structure came around in the 1960s, hemoglobin myoglobin other structures. Consider whether you are a scientist or crystallographer working in 1960.

Do you know how computational biology would be becoming a very, very critical contribution to your field? No, right. That is what happened. So, it is not always people who do not do their work correctly. But if you look at a coordinate file, if you look at the MTC file, they are huge files. So when people have the electron density, I am talking about x-ray crystallography because that contributes to nearly 90% of the structures and as a crystallographer.

I know when you are, you know, at the last stage of solving time, you get the answer of why you are doing that. You might ignore the proper building up of each atomic. But after 2000, when computers are becoming so essential in biology, and they are doing an amazing job, we gradually find that the ignorance of the extra crystallographers is affecting computational studies. If you are taking a PDB file and have no idea about the actual technique, consider this a 100% correct structure.

Unfortunately, this is not the case. So and I will keep discussing what I discussed before, so whenever this is my sincere request to all of you working in this field research purpose, especially if you have to perform structure repair.

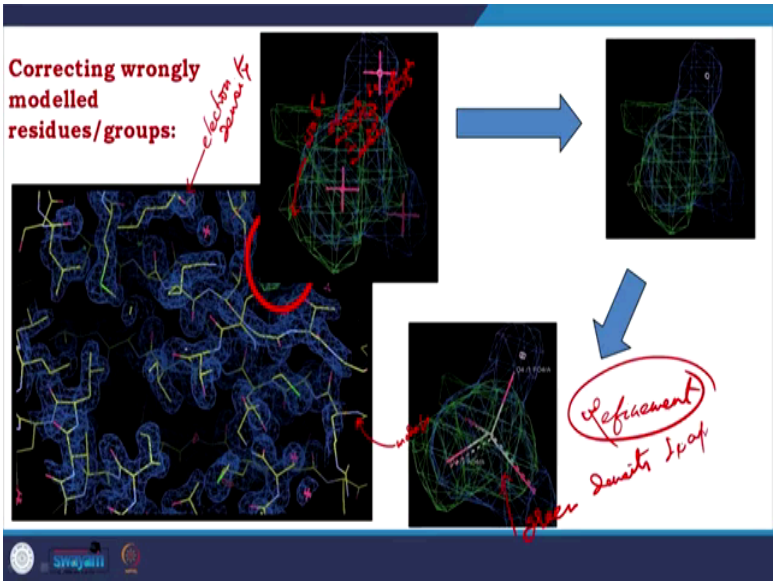
(Refer Slide Time: 16:20)



You have to identify missing residues and missing atoms. You have to correct the wrongly modeled residues or groups i and f if they are present in the structures you want to work. You have to take care of the protonation states of the different amino acids because when it is structured in a crystallo condition, what you are trying to develop in your system might be different. So, you have to take care of that. Unfortunately, you have to correct the rotamer confirmations like in crystallographic structures.

There are multiple rotamers, but in the represented PDB structure going for simulation, you have to provide all the one confirmation, so you have to take care of the steady classes because if somehow there are study classes, your program will not run. Here are 2 examples to show you because we are not doing any practical here.

(Refer Slide Time: 17:30)

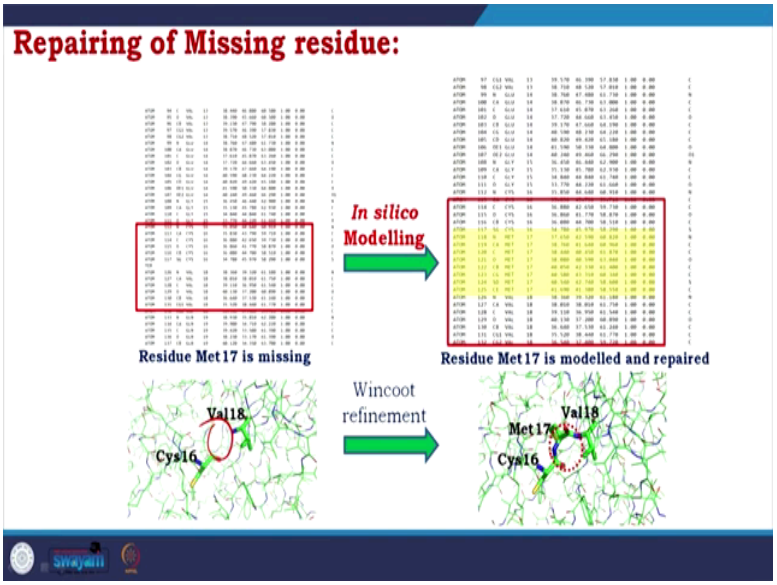


But you have to do hands-on to understand it better. This is the country presentation that we discussed in the last module. You see the models. The yellow-blue color is the model you can see. The clouds are called electron density, you already know. So, these are electron densities. These are models. These are water. Now, when you see green like that, you understand that some things should be modeled that are not modeled.

So, to give you an idea, and as I have already discussed in the visualization class, we have made these color combinations like traffic lights. So, when you see a green light, that means you should go. You should do something like in the traffic light you are allowed to go when you get red light that means you are not allowed to go in structure solution, you have put some atom in that density, where you should not put it, and you will always the in between.

So, I zoom in and see water molecules when I have taken care. We have cleaned the water molecule, and we have successfully modeled f-phosphate. When you run a refinement, you will see these green densities disappear.

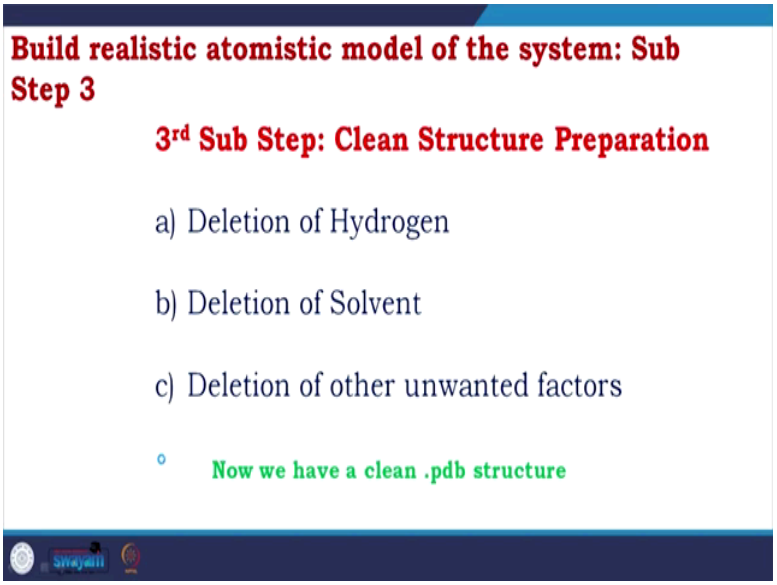
(Refer Slide Time: 19:54)



Another very critical thing is repairing missing residues in many cases. He says because of experimental restrictions problems, you could not model some residue, but when you are going for simulation, you have to model it. So, if you see here an example, the residue Met 17 is missing it you using in-silico modeling, we have modeled it repaired that and now, we take help up quote you know, in the program code it is not only a visualize and program as I always say, it is also a program by which you could do real experiments.

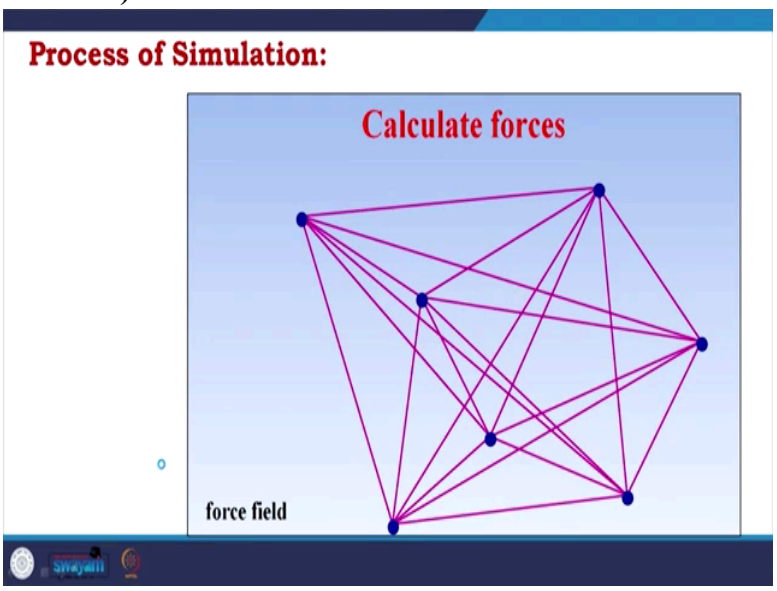
So, what we did, we did refinement, but refinement means, I have already explained many times, refinement is when you have those atoms you made a change. So, based on the already established library, you push it towards energy minimization. That is called refinement. So, you see that Methane 17 appeared but did not appear here. So this is called repairing missing residue. In such a way, there is a lot of repair, and it is always important to take a very careful look at your structures so that the initial structure you are making makes it flawless.

(Refer Slide Time: 21:38)



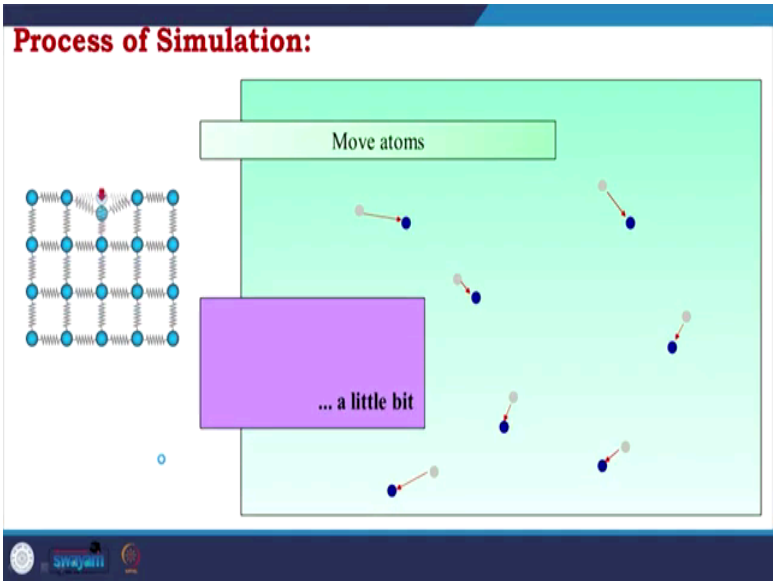
The third step is cleaning structure preparation, and you delete hydrogens, you delete solvents, you delete other things probably coming from like in crystallography, it comes from crystallography, so you check it out, you just delete them out. Now, you have a clean dot PDB structure. So, you have a realistic atomistic model of the system ready to go for simulation.

(Refer Slide Time: 22:15)



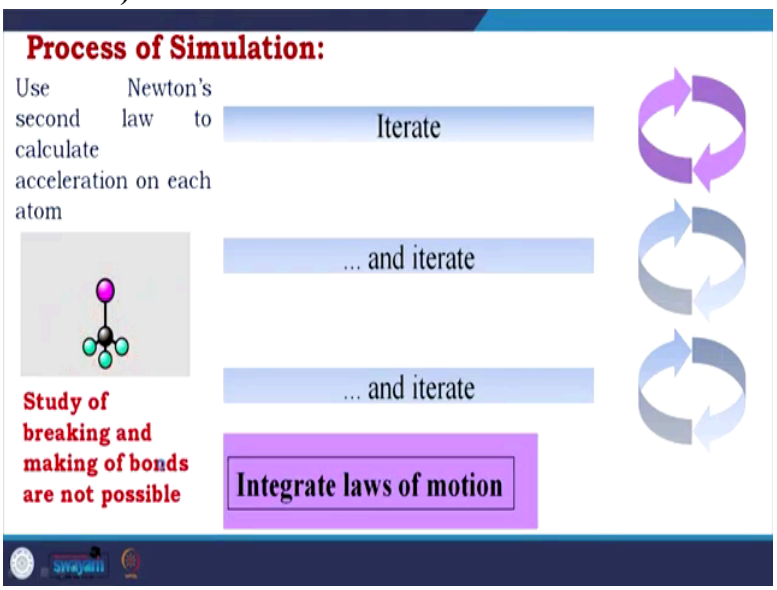
The second step is to simulate the behavior. So, what you have to do in the process of simulation, you have to calculate the forces in a force field.

(Refer Slide Time: 22:23)



So, what you have to do is you have the model, you have to keep the model in the force field, and you have to put force so that the atoms move from one point to another point, and if you see, there is a spring model, which is showing you how the force could be deforming that structure.

(Refer Slide Time: 22:59)

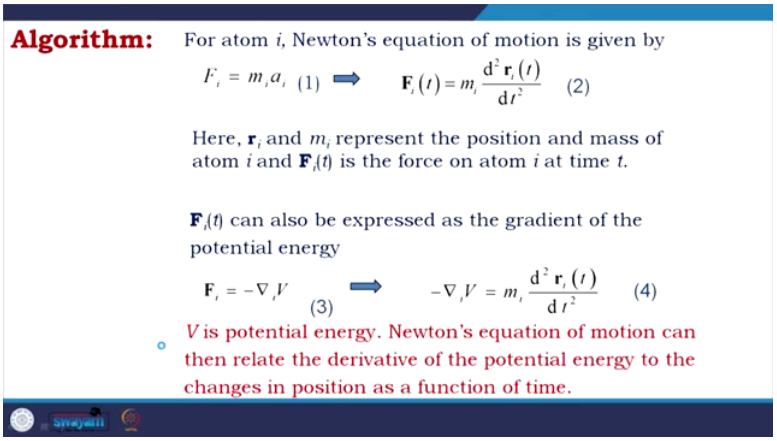


So, now, you do that, and this process you iterate, when you have anything it was creating this needs a computer algorithm to perform that iteration. You integrate with the laws of motion, so here we use Newton's second law to calculate the acceleration on each atom, and if you look at you are allowing the stretching by considering classical physics. Still, you are

not considering the bond breaking. So, the study of breaking and making bonds is impossible in this scenario.

Then, now, you should have a question. So, how I would study enzyme when it is taking care of a reaction, biochemical reaction classical MD simulation as I told is obeying the Newtonian physics. It is all only about the movement. Anything else you need to do? You have to set up specialized treatments. We will talk about that.

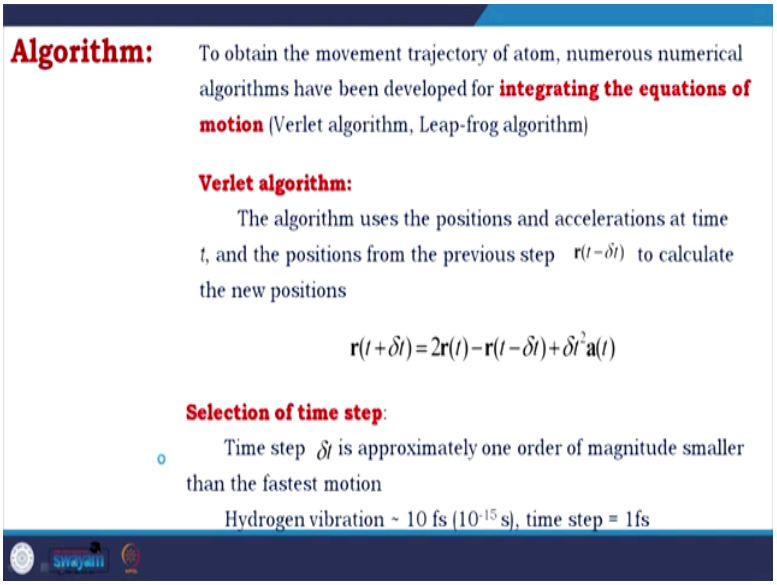
(Refer Slide Time: 24:32)



So, coming to the algorithm from atom i, Newton's equation of motion is given by F i = m a where m is the mass, and a is the acceleration, with the time, it is d 2 r r is the distance dt 2. Here r i and m i represent the position and mass of the atom i, and F i t is the force on the atom at time t. This acceleration could also be expressed as the potential energy gradient, F i = -del i V, so del i V = m d 2 r i t dt 2.

That means the potential energy V, with the help of Newton's equation of motion, can then relate the derivative of the potential energy to the changes in position as a function of time. Now, if you see you are in a situation where you put a poor, there would be a change and measure the change with respect to time, and when you put a force like you have a ball, you throw a ball, it would go, and it would come back. So, it would create a trajectory.

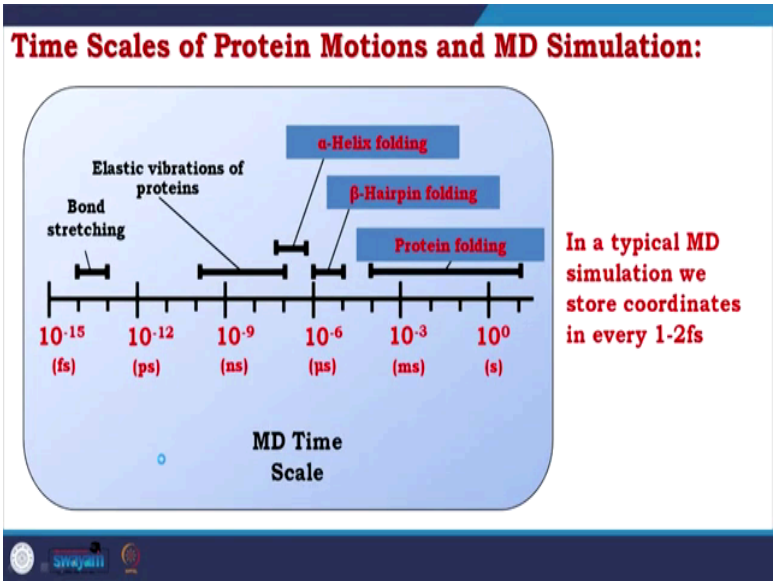
(Refer Slide Time: 26:32)



We have to calculate the trajectory to obtain the movement of the atom's trajectory. Numerous numerical algorithms have been developed for integrating the equations of motion. There is a Verlet algorithm, a fundamental and the first algorithm in this field work that velocity Leap-frog; there are many algorithms. So, as it is an introductory course, I will give an example of the Verlet algorithm.

The algorithm uses the positions and accelerations at time t, and the position from the previous step is t -del t to calculate the new positions showing r t + del t = 2 r t -2 r t -del t + del t 2 a t. So, you have a current position to get the next one, and you have to calculate the change from the previous one. Now, this would also help you to select the time step. The time step delta t is approximately one order of magnitude smaller than the fastest motion. So, for hydrogen, if you consider the diatomic molecular spring, it is 10 femtoseconds. So, the time step is 1/10 of that its1femtosecond.

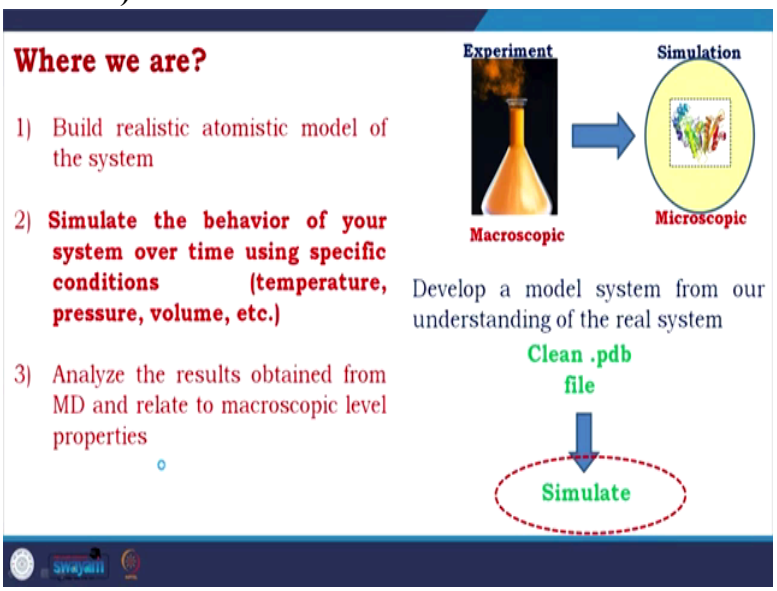
(Refer Slide Time: 27:32)



So, remember I was talking about this is the time scale of protein motion where bond stretching is around femtoseconds which I talked about, the elastic vibration is around nanoseconds, the alpha helix folding is a nanosecond microsecond, the beta-hairpin folding in the level of the microsecond and protein folding in the level of a millisecond to second when we perform as I discussed in detail in the previous class.

We are taking care of the unfolding of the protein, and we mostly want to see the molecular interaction. What we do we store every coordinate after 1 or 2 femtoseconds so that we can use them for minute analysis after the MD simulation is performed.

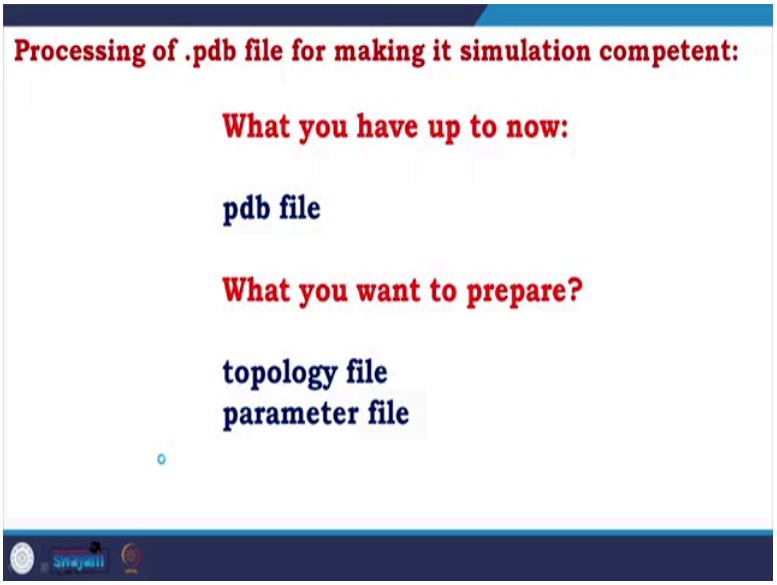
(Refer Slide Time: 28:28)



So, as I talked about the experiments where we are, we want to simulate the behavior of what we already have to build the system's realistic atomistic model. Now, we want to simulate the behavior of the system. So, we developed a model system from our understanding of the real

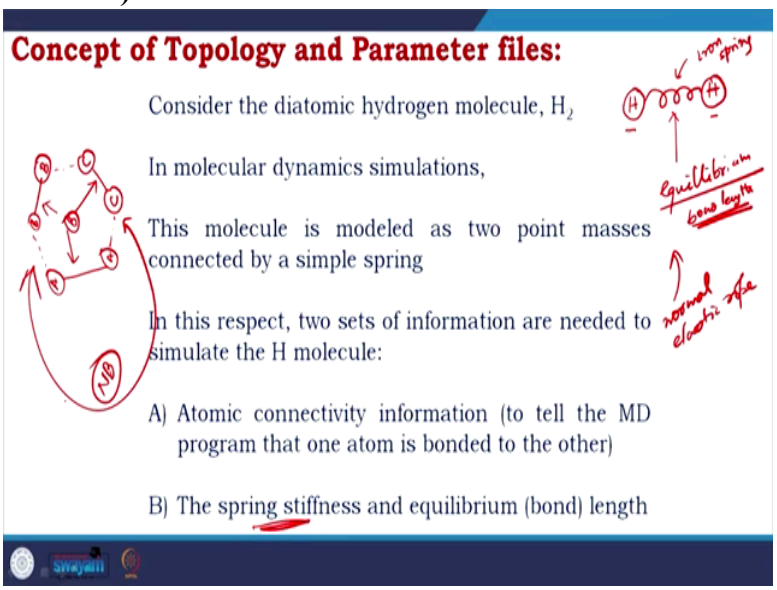
system and had a clean PDB file. Now, we think we use that clean PDB file to simulate, but that is not real.

(Refer Slide Time: 29:01)



Because you have the PDB file, you could not use it for simulation, so what you have now is the PDB file. What you want to prepare is a topology file and a parameter file. So, we have already started the simulation part, but in this part, we will not do the simulation at the first step. We will prepare the file properly so it can be under simulation.

(Refer Slide Time: 29:34)

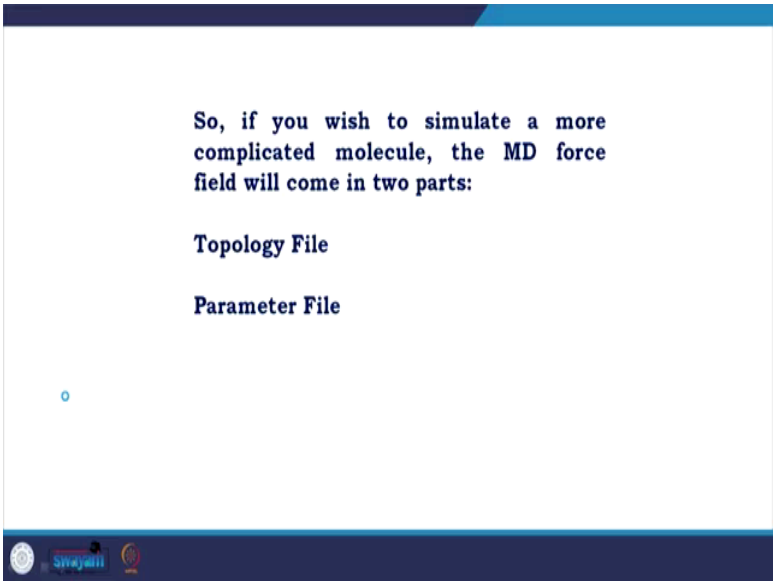


So first, we will consider the concept of topology and parameter files. To do that, as we have already told you to, consider a diatomic hydrogen molecule. In molecular dynamic simulation, this molecule is modeled as 2-point masses. You could see 2 point masses connected by a simple spring. In this respect, 2 sets of information are needed to simulate the hydrogen molecule. What are they?

One atomic connectivity information. So, remember I told you this spring that if you put a force, it would be expanded, it could be contracted, and then it comes to the equilibrium position. This equilibrium position is the bond length that we will talk about the connectivity when you put a force, and it cannot go beyond the distance. It could not come lesser than a distance. That is what we get here. Atomic connectivity is another of the spring's stiffness and equilibrium.

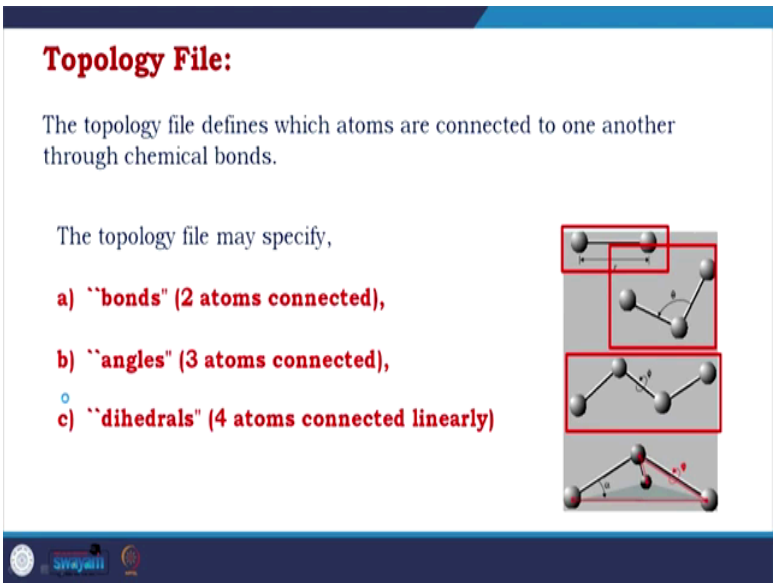
So, suppose I am using an iron spring whereas you are using a normal elastic rope, so they would have their stiffness variable that could be that 2 factors. One could tell the MD program that one atom is bonded to another because if you look at a system like this, where this is bonded, this is non-bonded, suppose A's are bonded, B's are bonded, and C's are bonded. So, there are some bonding information and some non-bonding information. We must tell the program that these are connected. You cannot put them miles apart. So that is atomic connectivity, and the second is the spring stiffness and equilibrium.

(Refer Slide Time: 32:15)



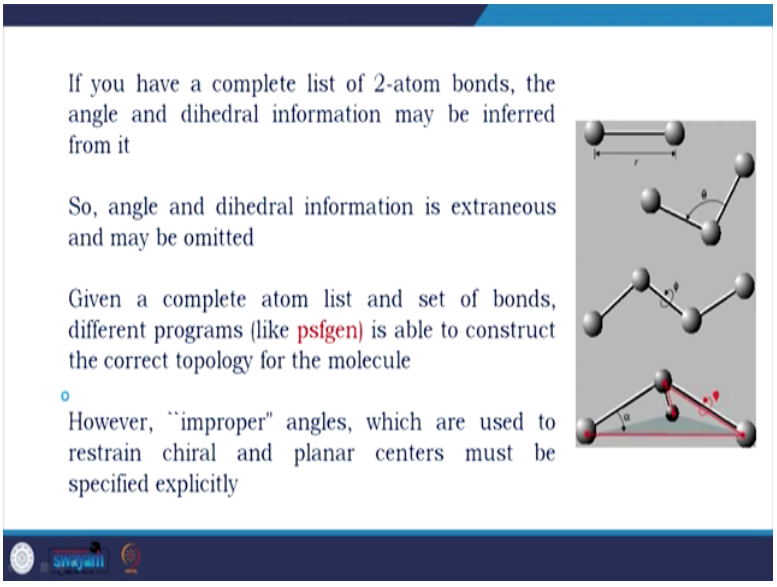
So, if we wish to simulate a more complicated molecule, the MD force field will come in 2 parts topology file parameter file.

(Refer Slide Time: 32:29)



The topology file defines which atoms are connected through chemical bonds. So, what it will talk about is bonds. We already know bonds mean 2 atoms connected angles, 3 atoms connected dihedrals, and 4 atoms connected for dihedral. There are 2 things proper and improper dihedral when we are talking about proper dihedral.

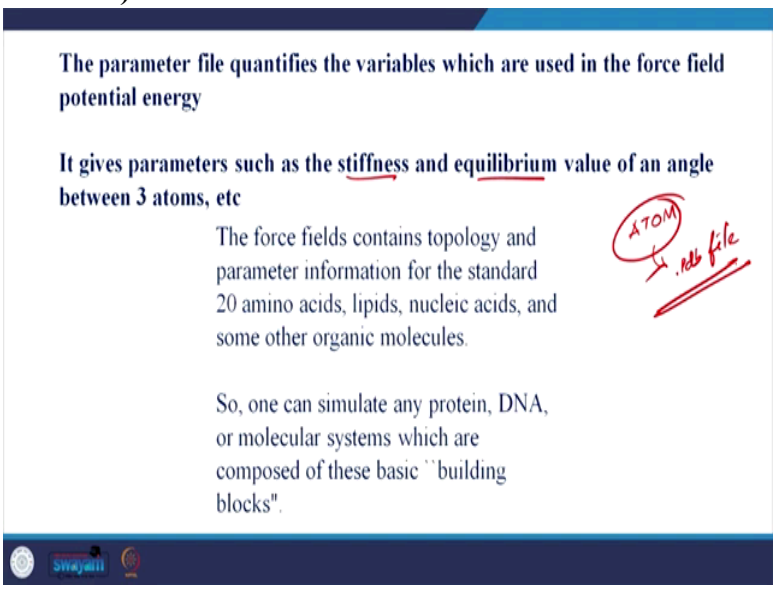
(Refer Slide Time: 33:03)



So, if you have a complete list of 2-atom bonds, the angle, and dihedral information may be inferred. So, if you have a library with bonding connectivity, your angles and I had droughts could be computed. So angle and dihedral information is extraneous and could be omitted from the library because it could be calculated with the already established connectivity information.

Given a complete atomic list and set of bonds, different programs like psfgen, a program from one simulation package in AMD, can construct the correct topology for the molecule. However, as I was talking about, improper angles must be specified explicitly, which are used to retain chiral and planar centers.

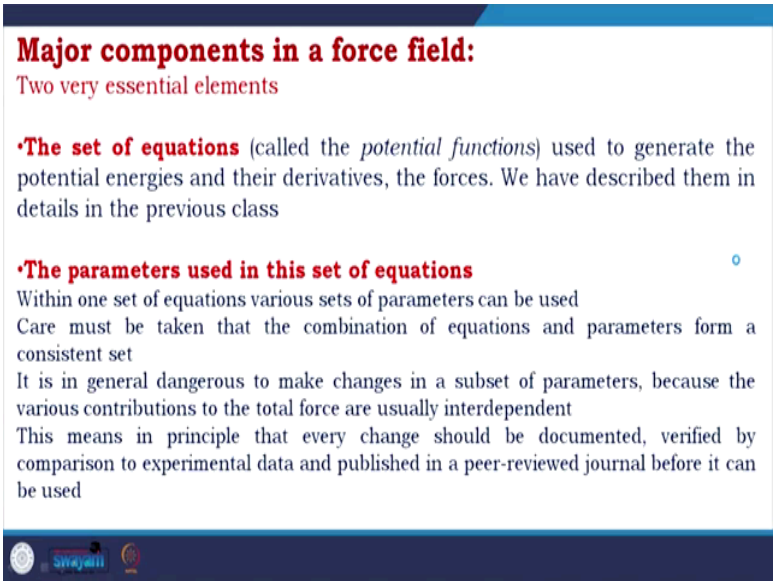
(Refer Slide Time: 34:13)



The parameter file quantifies the variables used in the force till potential energy. It gives parameters such as the stiffness and equilibrium parameters, an equilibrium value of an angle between 3 atoms, etc. Now, this could differentiate the force fields, the force field contents, topology, and parameter information for the standard 20 amino acids, lipids, nucleic acids, and some other organic molecules.

Remember, when discussing the PDB file, I have shown you the atom person and a hetero atom portion. The atom portion in the dot PDB file is known information should these be in the force field. Using them, one can simulate any protein, DNA, molecule, or system composed of these basic building blocks. So, all the biological macromolecules coming from protein DNA and RNA could be simulated.

(Refer Slide Time: 35:37)

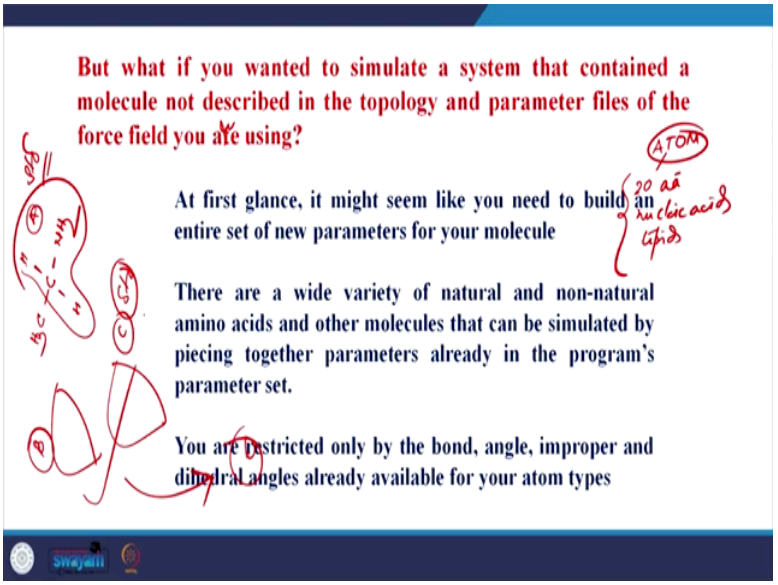


And these parameters when you go to theory, these parameters have different changes and different contributors, which gives us different force fields. There are 2 very essential elements in the force fields. On the set of equations, they have called the potential functions used to generate the potential energies and their derivatives the forces. Remember, we discussed the components and their equations, the stretching, bending dihedral, Vanderwaal, and Columbic hydrogen bonds. They are the same.

But their contribution factors constants are changing according to the theory people have developed, and there are the parameters used in this set of equations. Within one set of equations, various sets of parameters could be used. What are they? We have to take a very close look at the individual force fields here. I am not going into that, but I am talking about the fact that a combination of equations and parameters must form a consistent set.

So, it would help if you had an equation and related parameters. Changing the sub-set of parameters on your own is generally dangerous because the various contributions to the total force are usually interdependent. So, you are changing one thing without understanding that this change could affect other calculations you could not. So, in principle, every change should be documented, verified by comparison to experimental data, and published in a peer-reviewed journal before it can be used.

(Refer Slide Time: 37:40)

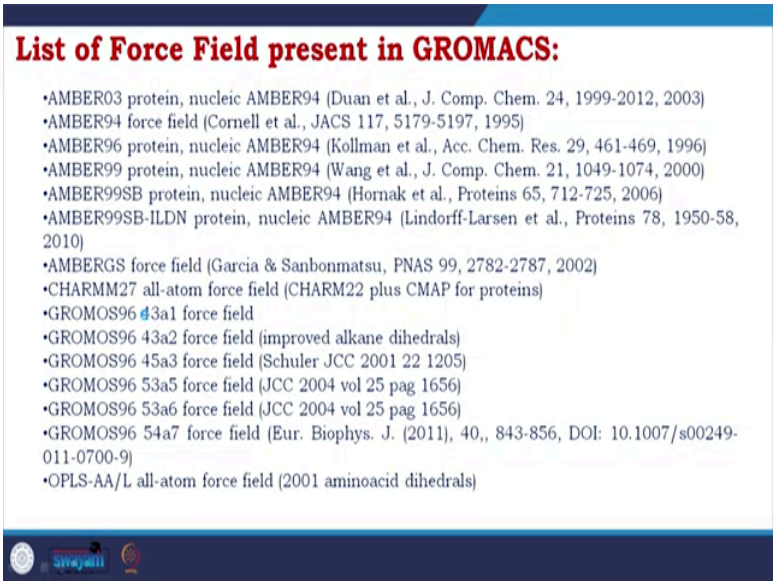


e

So, I talked about how you wanted to simulate a system containing a molecule not described in the topology and parameter file of the force field you are using. Do you understand? So before I tell you to have the atom data in the PDB, about 20 amino acids, nucleic acids, lipids, and all. But if someone or some system some molecule does not have that, it seems that you have to build the entire set of parameters if they are not matching, but that is not the actuality and is a very relieving factor.

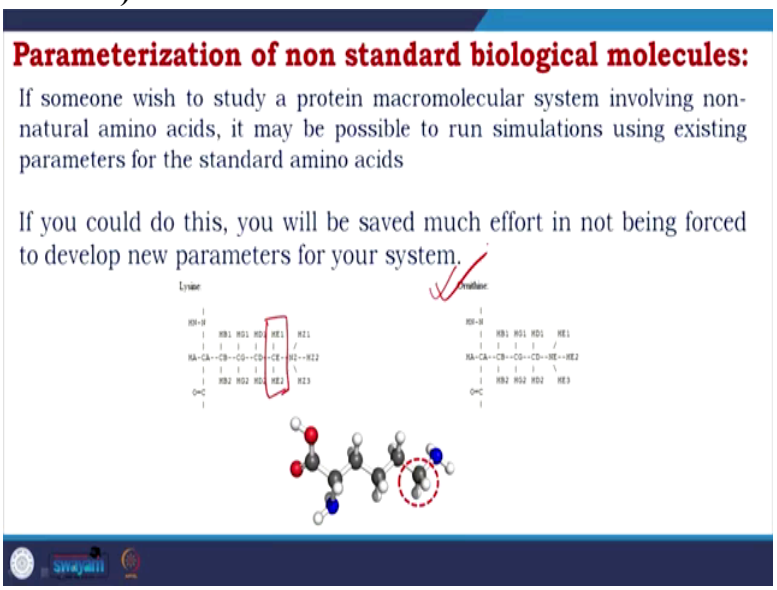
A wide variety of natural and non-natural amino acids and other molecules can be simulated by piecing together parameters already in a program's parameter set. This is very interesting. Suppose there is suppose you need this part from one molecule and the rest of the part from B molecule. These are all standard A, standard B Standard C. By having parts from these, these, these, you get molecule D, whose parameter was not there. This is also helping you. So you are restricted only by the bond angle improper and dihedral angles already available. Not by re-conformers of how they formed.

(Refer Slide Time: 40:03)



This is the list of force fields present in GROMACS. I take gromacs as an example. As I told you in this introductory course, I am just talking about showing you the force field. Hopefully, we will get the opportunity to take a course in the future, and I will describe all the program packages that force fields and all. So, you will see that there are majorly AMBER, CHARMM, GROMOS, and OPLSA. It is difficult to explain how they work, but if you ask me how you should select them, I would say one go through extensive literature searching to understand what is good for your protein and then do trial and error.

(Refer Slide Time: 40:55)

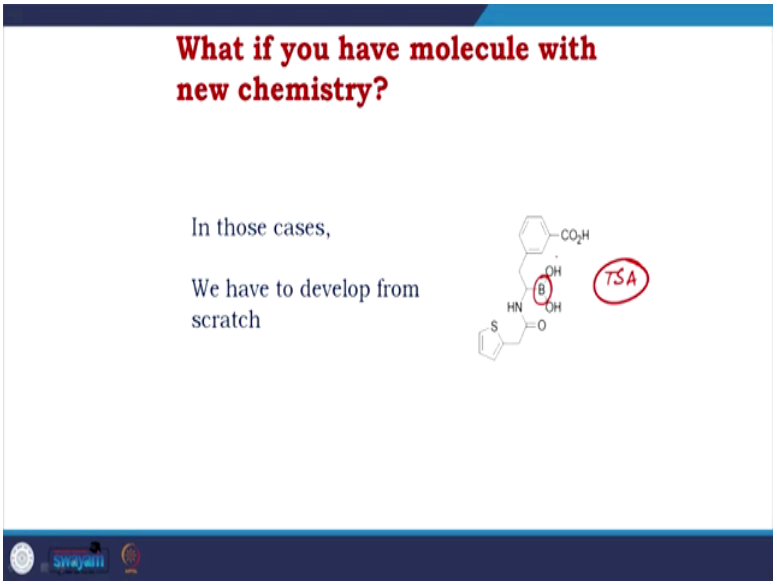


So, I talked about this, if someone was to study a protein macromolecular system involving a non-natural amino acid, it might be possible to run the simulation using existing parameters for the standard amino acids. If you could do this, you will be saved much effort by not being forced to develop new parameters for your system. So, I have taken an example you know

about lysine, a natural amino acid, and ornithine non-natural amino acid. If you compare, you will see these parts are the extra part in lysine.

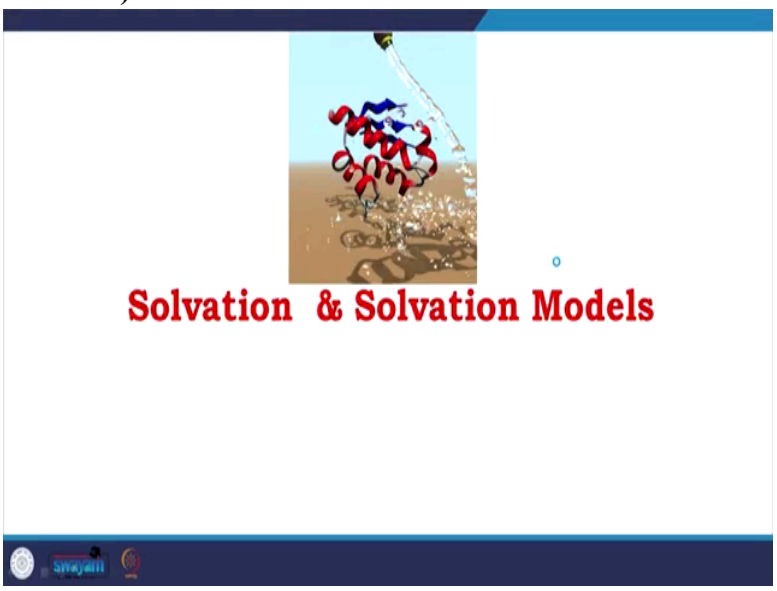
So, if you remove that part's parameter, you use the rest of the parameter for ornithine. You do not need to develop an entirely new system of ornithine. This is an easy system to demonstrate, but as I told you in the earlier slide, you could combine pieces from very different systems and apply them to a successful simulation run.

(Refer Slide Time: 42:07)

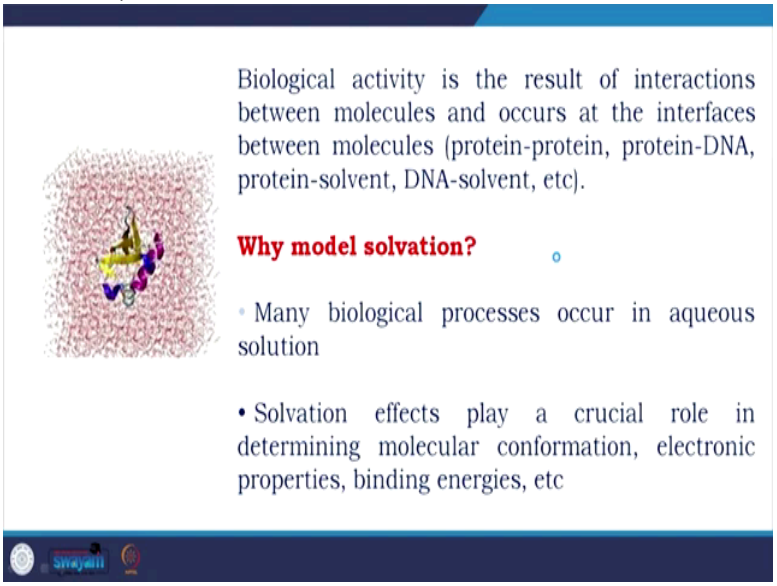


So, what if you have a molecule with entirely new chemistry? In those cases, we have to develop from scratch just like the boron compounds, as you see in the literature, are becoming popular, especially as a transition state analogue mimicking compounds. So, boron chemistry is not there in the parameters. So you have to develop your own.

(Refer Slide Time: 42:38)

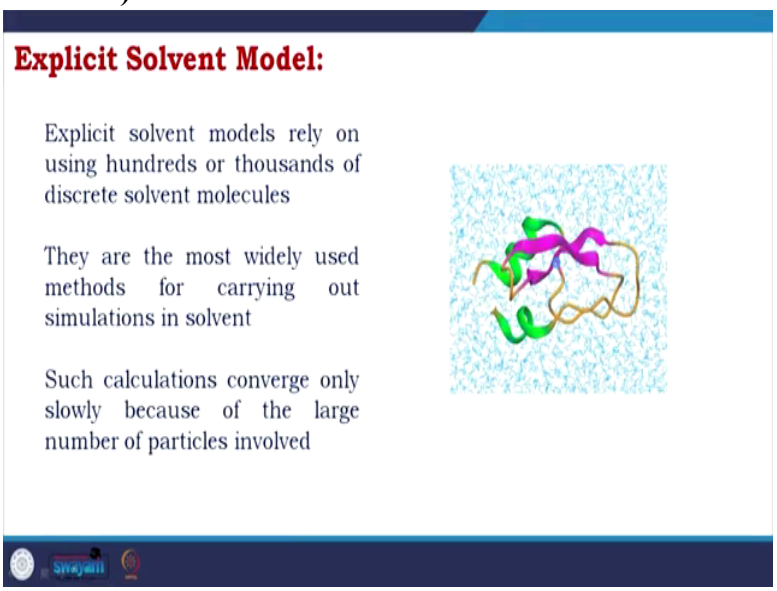


(Refer Slide Time: 42:42)



You are coming to salvation and salvation models. Biological activity results from the interaction between molecules and occurs at the interface between molecules, protein-protein, protein DNA, protein solving DNA solving. So, these are the normal system we are looking for. Most of the biological processes occur in aqueous solution. If you look at the cytosol, it is mostly accurate. So that is why the biological macromolecules interact in aqueous solution. Solvation effects are crucial in determining molecular confirmations, electronic properties, binding energies, etc.

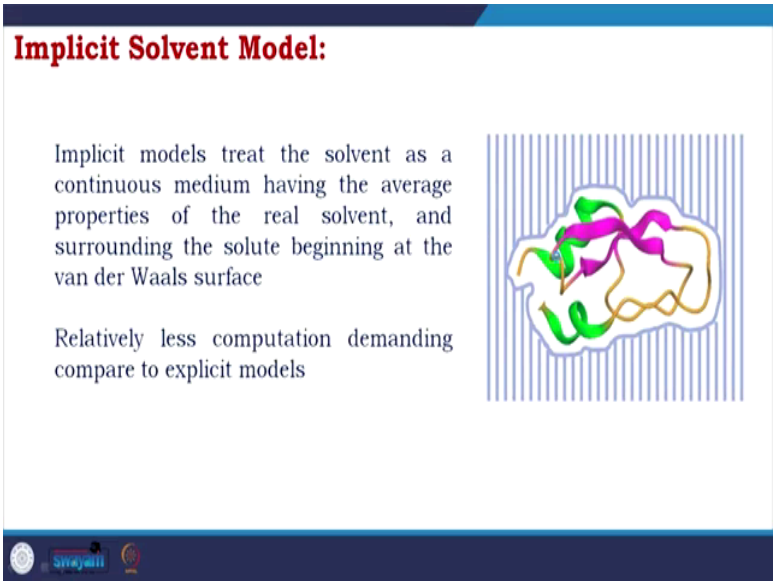
(Refer Slide Time: 43:22)



There are 2 types of models explicit solvent model and implicit solvent model. Explicit solvent models are real models like you have 100s and 1000s of water molecules or solvent molecules. Their explicit solvent model is the most widely used method for simulation in the

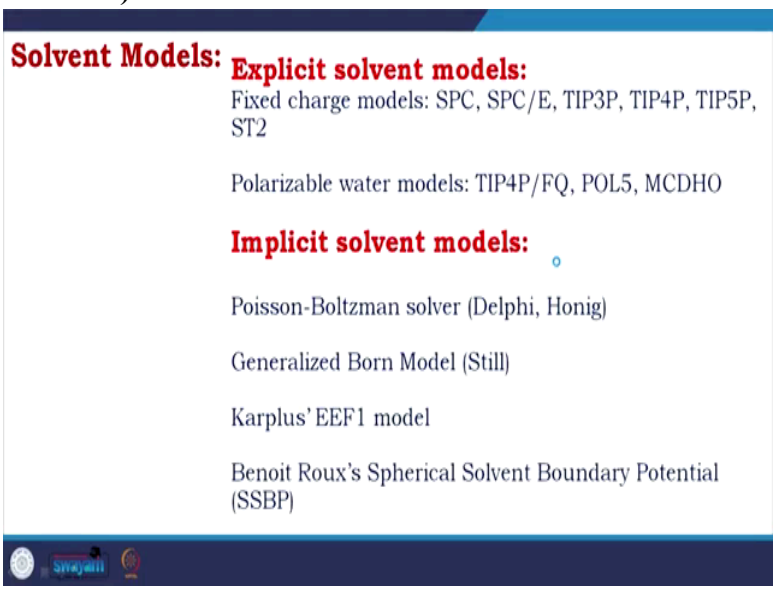
solvent because they are realistic, but because they are many numbers of computationally expensive.

(Refer Slide Time: 43:51)



The implicit solvent is a model. It treats solvent as a continuous medium and the average properties of the real solvent, so it is a calculated thing surrounding the solute beginning at the Van der Waals surface. As you have seen here because this is already a calculated model, the computational complexity is much less, so much less time is required.

(Refer Slide Time: 44:22)

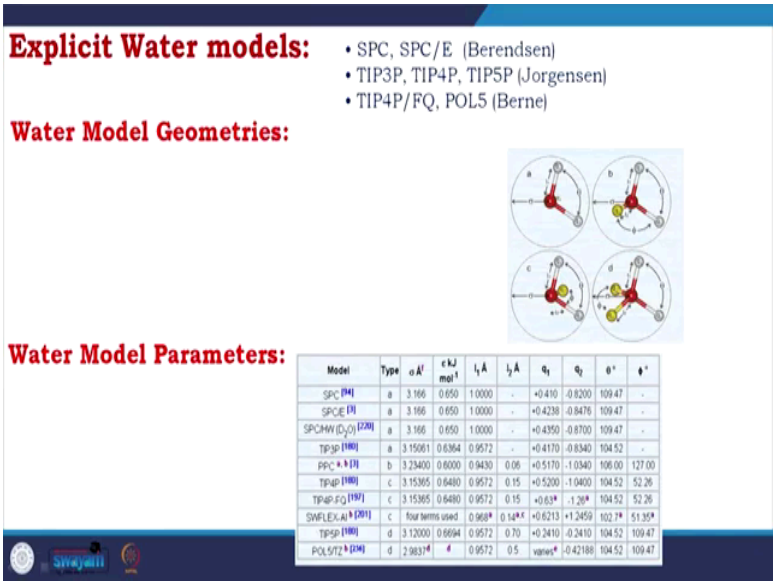


Examples of the explicit solvent model. There are 2 types of explicit solvent models fixed charge model, where the solvents are charged. Still, it is not considered that this charge influenced others like SPC SPCE, TIP3P, TIP5P, and ST2, and all polarizable water models are charged models are where they are charged. Still, they influence others TIP4P FQ, POL5,

and MCDHO are examples of implicit solvent models. As I said, they generally come from the theory's development.

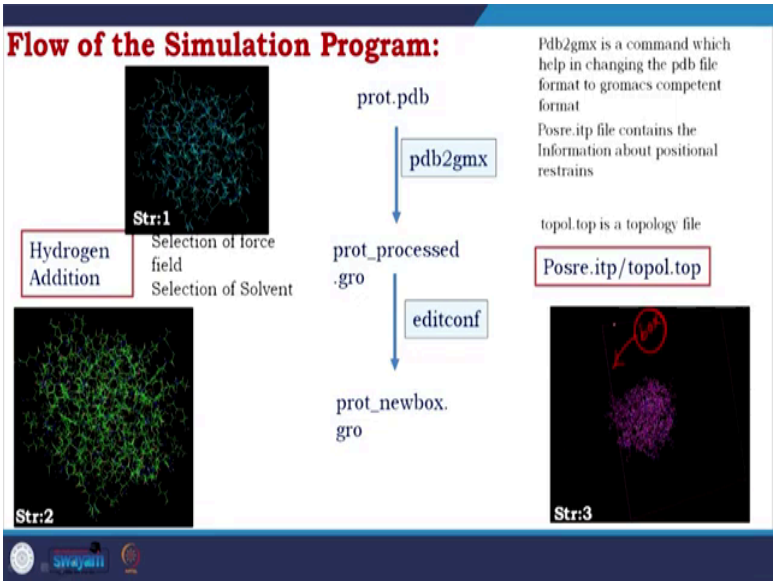
So Poisson-Boltzmann solver, which is under the program Delphi by Barry Honig, generalized born models still Karplus, you know Karplus, he is the, we could say the heart of MD simulation, and it is further proceedings he is a Nobel laureate, EEF1 model, and Benoit Roux from the University of Chicago spherical solvent boundary potential SSBP model.

(Refer Slide Time: 45:34)



Suppose you look at how the differentiation happened. In that case, this is the water model geometry, where if you see you will see that they are considered the distances the charge, the charge is fixed or polarizable, and considering those many factors, they are included in the model and depending on the differentiation, like sigma is depending on the solvent surface charge, if Cylon is dielectric constant the distances the charges the angles, depending on them, you see, there are differences in the model.

(Refer Slide Time: 46:18)



So, coming to the flow-up simulation program, we already got the clean structure dot PDB; I named it Structure 1. So, the file is prot dot PDB which is protein dot PDB. Protein dot PDB converted to proc underscore process dot gro by the help of a common. So, I am taking the gromacs program package as an example. So, PDB to genomics is a command that converts the PDB formatted one to gromacs competent.

Here we add hydrogen we select the force field. I talked about the force field of around 15 force fields. I have shown you have to select the force field here, and you have to select the solvent model like TIP3P, TIP4P, SPC any model you will use. So, that PDB to gmx common gives brought underscore process dot gro the gro file is a gromacs competent file with all the information this is the main file associated with posre dot itp and topol dot top.

Posre dot itp is a very interesting and very critical file for the simulation to run as I talked about, your atomic connectivity information of the covalent bond already told which atoms are being restricted to which length but for non-covalent interaction, there is no limit to a loop an atomic loop could you know be here and it comes it could you know even travel a huge distance, but, at the same time, you cannot like to consider infinity as your simulation system.

So, you put the restriction posre dot itp file content, the information about the positional risk friends, which are already given here, topol dot top is a topology file. You see, from structure 1 2 structure to the hydrogens are added, you might ask me a question in many structures. We are downloading hydrogens from PDB, so why have you taken the initiative separately, or

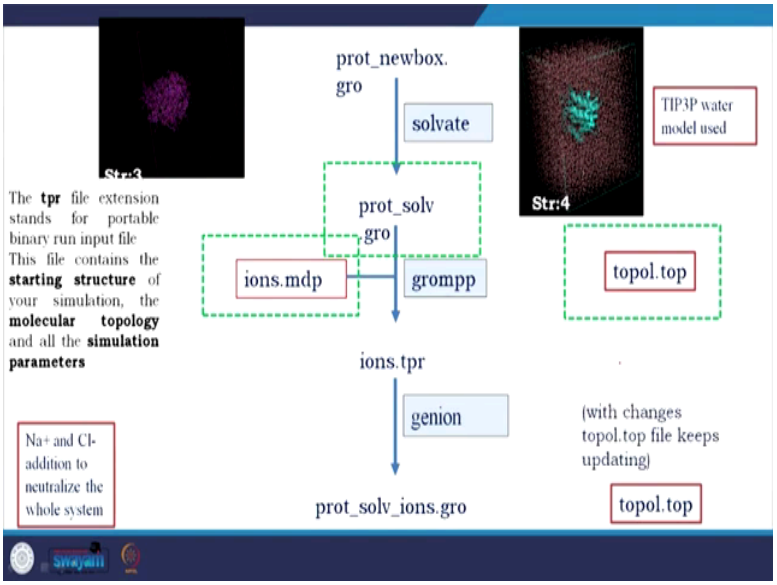
why is the program adding it separately? There are many reasons, but I must tell you the most important one.

If you look at among the 3 high-resolution structural biology techniques, X-ray crystallography, NMR, and cryo-electron microscopy, up to now, 90% of the structure is solved by X-ray, and as we have already gone through the detail of the X-ray crystallography process, you now know very well X-ray could not deflect hydrogen. So, the protein structures, the 90% contributed structures, are not having the information of hydrogen. I am surprised, yes.

Then from where the hydrogens are coming, the hydrogens are coming through theoretical calculation. So, different software has different principles of calculations before simulation. We delete them and use a unified program to add hydrogen, which you see here; the prot dot PDB had no hydrogen prot underscore processed dot gro had hydrogen. Now, you come to the next one where you have prot underscore new box dot gro with a comment editconf.

Editconf is taken care of conformation where it helps you to develop if you look at structure 3, a box should create a box, and the box is created in such a way if you consider a sphere around your protein, you will see equidistance maintenance of the box. So, in the box, the protein is kept in a centralized position.

(Refer Slide Time: 51:24)



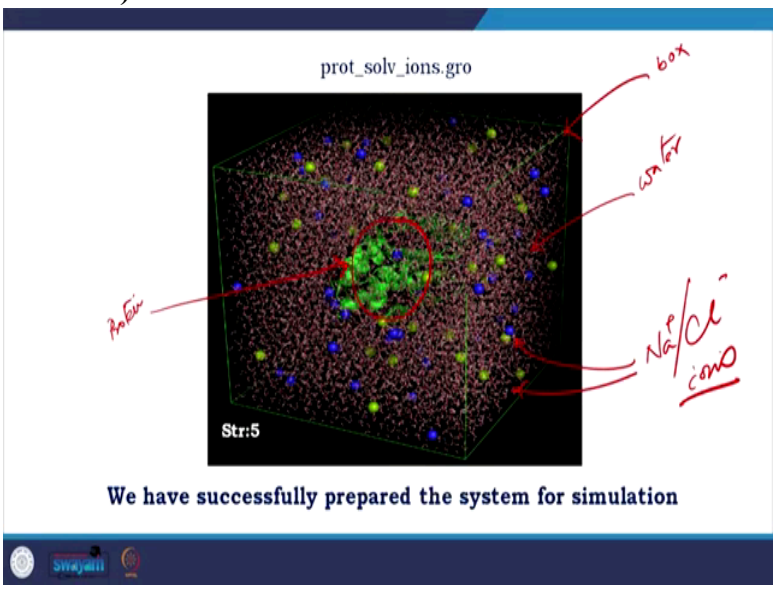
In the next step, protein underscores new box dot gro is transferred to protein underscore solv dot gro by solvating it, so we have already selected the solvent. From structure 3 to structure

4, your protein is solvated with water molecules. So, in this case, we have selected the TIP3P water model. Topol dot top would be an associated file. Now with this prot underscore solv dot gro, ions dot mdp is added dot mdp is a molecular dynamics parameter file by associating 3 files ions dot mdp prot underscore solv dot gro and topol dot top.

They make the file ions dot tpr by the common grompp. Grompp would be a comment that could come when you merge files, especially mdp dot gro and dot top dot gro. As I told you before is a file that is gromacs competent. The dot tpr file is the portable binary run input file extension. This file contains your simulation's starting structure, the molecular topology, and all the simulation parameters.

So that is why it marks everything and brings it into a dot tpr file, and dot tpr file, with the help of the common genion generation of ions, comes into prot underscore solv underscore ions dot gro here. Remember I told you we are mimicking the in-vitro condition? So, we are developing a buffer, and we add sodium and chloride to neutralize the whole system. Topol dot top would be updated as I said with changes topol dot top file keeps updating continuously.

(Refer Slide Time: 53:54)



So, this is the prot underscore solv underscore ions dot gro file. If you look at this file, you have the protein. This file has water; these files have the there is a box there are ions. So now we have successfully prepared the system for simulation. So yes, we need the PDB file, but the PDB file has to be processed properly to come to this dot gro file. So today, we have

started with understanding the force field converting to the algorithm, building a realistic, atomistic model, and further processing the model towards simulation.

We will perform the rest of the simulation process in the next class. Thank you very much for listening.