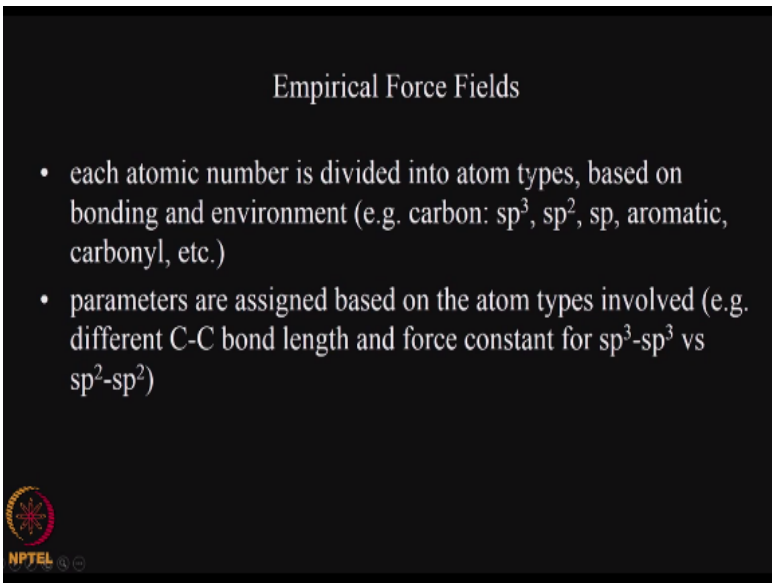
Computer Aided Drug Design
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture - 17
Molecular Mechanics / Force Field

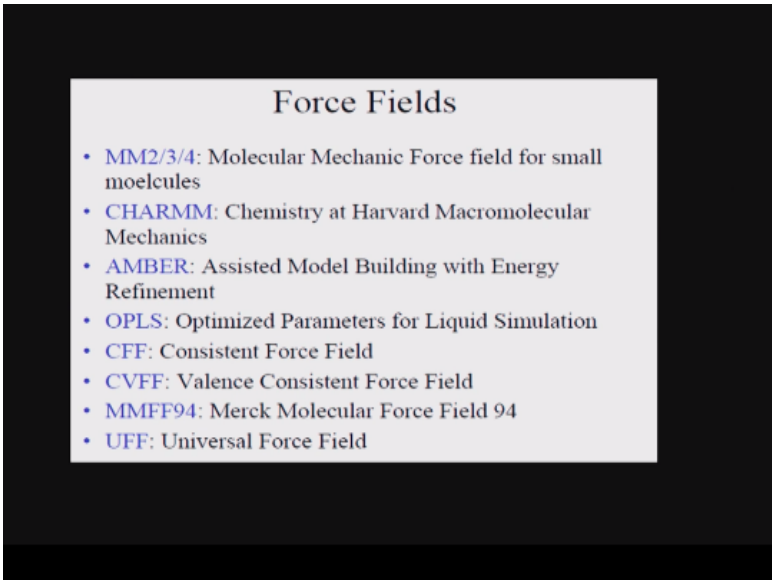
Hello everyone, welcome to the course on computer aided drug design. We will continue on the topic of force fields or molecular mechanics, this is going to be a very important topic as I mentioned so we have to spend lot of time on this, now these empirical force fields each atomic number is divided in to atom types, okay, based on the bonding and environment.

(Refer Slide Time: 00:36)



For example, carbon can be many types, sp3 carbon, sp2 carbon, sp carbon, aromatic carbon, carbonyl carbon okay then cyclopropane type of carbon, cyclopropene type of carbon so on actually okay, and then parameters are assigned based on the atom types, okay, so each of these carbon may have different carbon-carbon bond length okay and the constant where we use that also will change depending upon sp3-sp3 or sp2-sp2 and so on. So we will look at some of those examples.

(Refer Slide Time: 01:08)

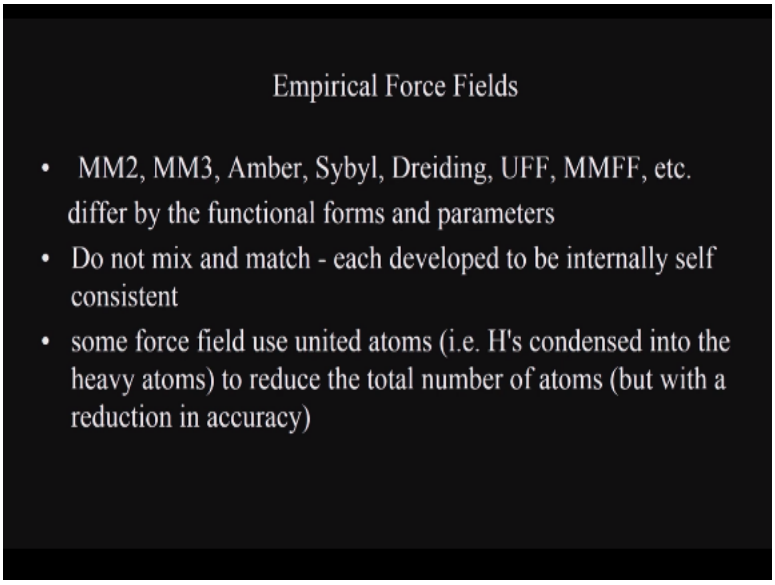


There are different types of force fields depending upon what type of parameters they choose, what type of equations they choose okay, like one family MM2, MM3, MM4, these are molecular mechanics force field for small molecules, these are very good, MM2 or MM3, is good enough for small molecules, if you are looking at drugs, okay. CHARMM: This is Chemistry at Harvard Macromolecular Mechanics, this is another force field.

AMBER: Assisted Model Building with Energy Refinement, OPLS: Optimized Parameters for Liquid Simulation, CFF: Consistent Force Field, so this is also used nowadays quite a lot, CVFF: Valance Consistent Force Field, then MMFF: Merck Molecular Force Field 94, then UFF: Universal Force Field, these are just few examples, but there are many many many types of force fields which are there.

And generally I think MM2, the CVFF are quite used very commonly and of course CHARMM and AMBER is also used quite commonly.

(Refer Slide Time: 02:20)



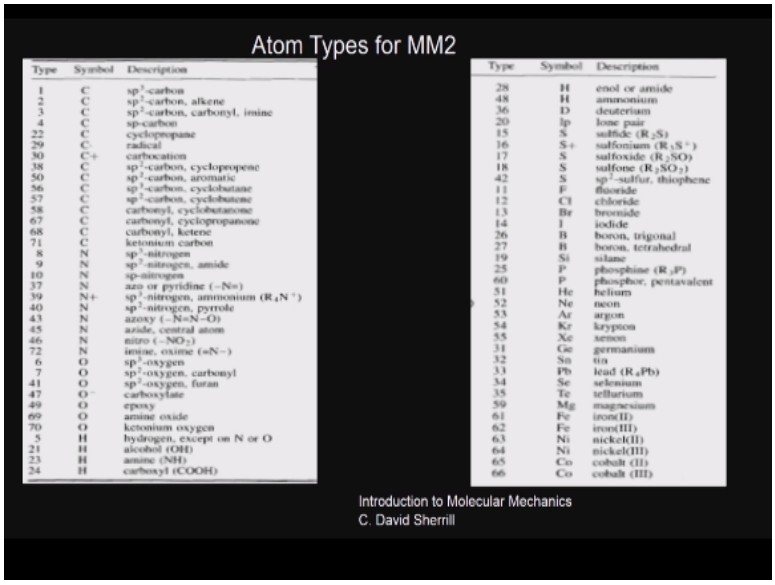
Okay so these force fields like MM2, MM3, AMBER, Sybyl, UFF, MMFF et cetera they differ in the functional forms what type of terms are used for bond stretching or angle bending and so on, okay, and the parameter values, parameter values will change so the functional forms maybe differing from one force field to another and the parameters mainly differing from one's field.

Do not mix and match, this is very, very important, so if I use one say MM3, I will do all my calculations for all the set of molecules with MM3, I will not mix and match because each set of force fields internally self-consistent, so remember that, like I said the functional forms within force fields may differ, the parameter values may differ, so if I use for some molecules, MM3 and some molecules I use Amber I will get different answers which is completely wrong.

Some force field use united atoms that mean what do they do is, they condense hydrogen H with the heavy atoms to reduce the total number of atoms, but there is some reduction in accuracy so instead of putting H separately, carbon separately they may have united atom which may condense H into the C and they may have parameter so, but little bit of inaccuracy but the number of calculations reduce dramatically.

So for example if I take MM2 let us take MM2, look at the different types of carbons MM2 has.

(Refer Slide Time: 03:51)



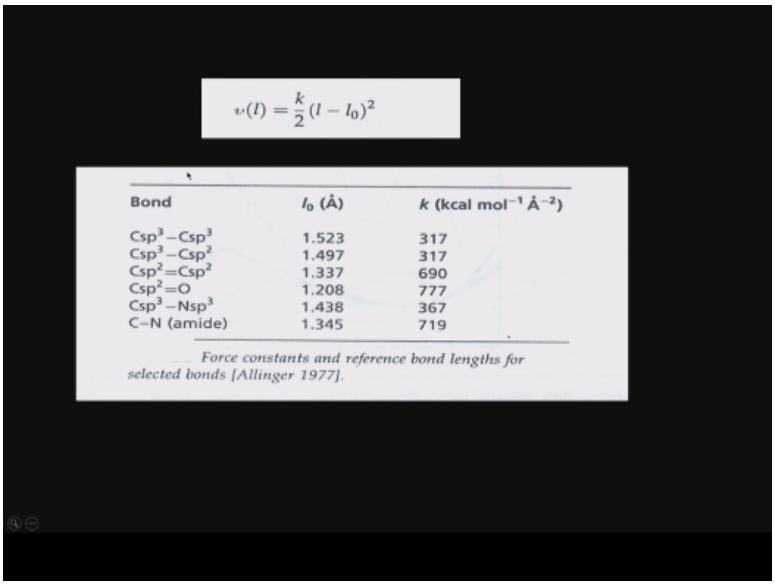
Okay, different types of carbons, sp3 carbon, sp2 carbon in alkene, sp2 carbon in carbonyl, imine, sp carbon, cyclopropane carbon, radical carbon, carbocation, sp2 carbon cyclopropene, sp2 carbon in aromatic, sp3 carbon in cyclobutane, so each of these carbon differs, so you may have different force constants, you may have different r not values and similarly look at this nitrogen, so many types of nitrogen okay.

Sp3 nitrogen, sp2 nitrogen, amide type of nitrogen, sp nitrogen, pyridine, ammonium okay then pyrrole, azoxy N double bond N and so on, nitro NO2 imine and similarly oxygen, so many types of oxygen, sp3 oxygen, oxygen carbonyl, oxygen furan, oxygen carboxylate, epoxy, amine oxide, ketonium oxygen so many different types of, so MM2 has large database for so many types of carbons, nitrogen, oxygen, hydrogen, sulfur, chlorine, bromine and so on including even metals as you can see here.

Okay so if you have cobalt or nickel or iron or magnesium, selenium, tin okay, so you are okay because MM2 has parameters for these so calculations for all these energies can be done properly, but if you have some other metal obviously it cannot handle, you may have to go for some other force field which can handle that particular metal, like for example okay, platinum maybe that is not in this list, okay, only some of them.

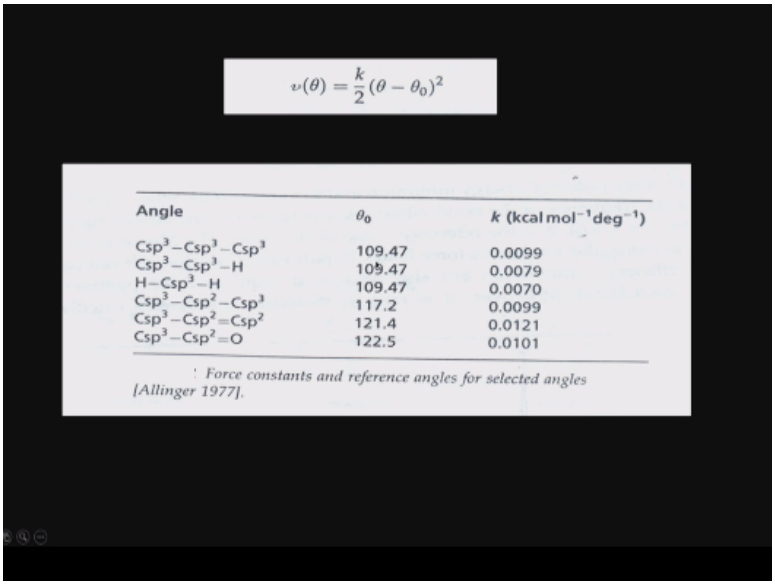
Sometimes some software will assume some other similar group metal and replace it with what they have, so each of these force fields will have different set of atom types so different parameter values so it is always good to use only one force field for all your calculation for example.

(Refer Slide Time: 06:12)



Let us look at this sp3 carbon, sp3 carbon, so the L not that is r not which we use there can be 1.523 whereas sp3, sp2 it becomes 1.497 and similarly the force constant also changes quite a lot depending upon the type of carbon right, sp3-sp3, sp3-sp2, sp2-sp2, sp20 that is carbonyl, sp3 N sp3 and so on actually, okay and similarly if for angle bending you have theta – theta0.

(Refer Slide Time: 06:50)

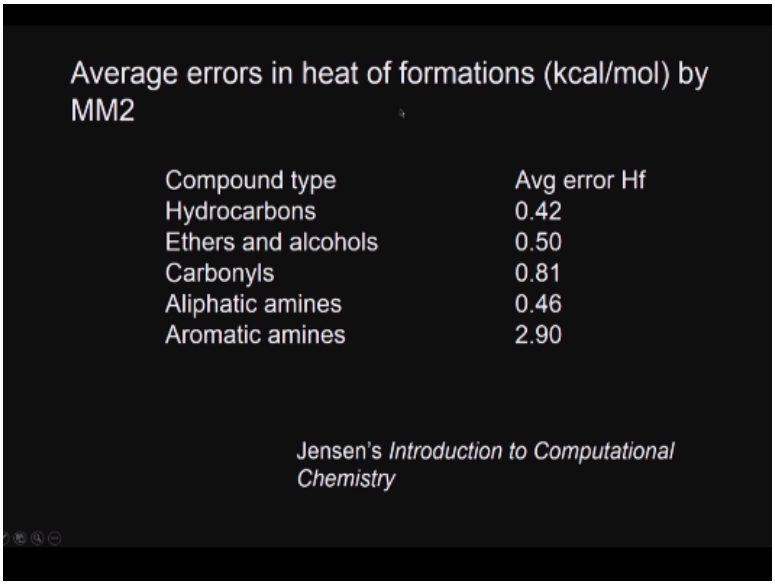


So we can have different theta0 values and different constant values here depending upon say for example, carbon-carbon carbon in sp3 carbon-carborn in sp3 with hydrogen H you get some other number here in force constant and so on actually. Okay, you see sp3, sp2, sp3 carbon the theta0 varies whereas this sp3, sp3 tetrahydral 109 okay, so these numbers will change, these numbers will change depending upon the atom type and the environment.

So remember so like that the database for any force field will be stored so when you draw a structure and you mention these are the different types of carbon, these are different types of oxygen I have automatically these numbers will be selected so the bond stretching energy, bond bending energy, torsion energy, nonbonding interactions will be calculated and the total energy of the molecule will be estimated.

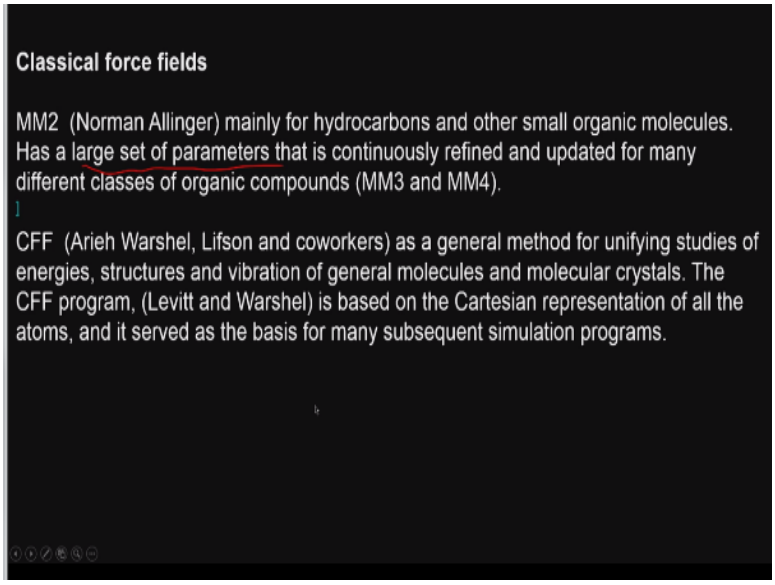
This is how these molecular mechanics force field approach is, so the more number of atom types I have it will try to cover as much as possible in the overall space okay, so that means your database becomes large so they will get these values you know either from experimental reported in literature or though ab initio report in literature, okay.

(Refer Slide Time: 08:15)



Okay so what is average error in heat of formation. When you calculate heat of formation using MM2, the average error on heat of formation may vary like this okay, so as you can see hydrocarbons region will be good whereas when you go to aromatic systems, amines it can be quite large, okay, so the errors depending upon how accurate you want your calculations to be you can decide and what you wanted to okay.

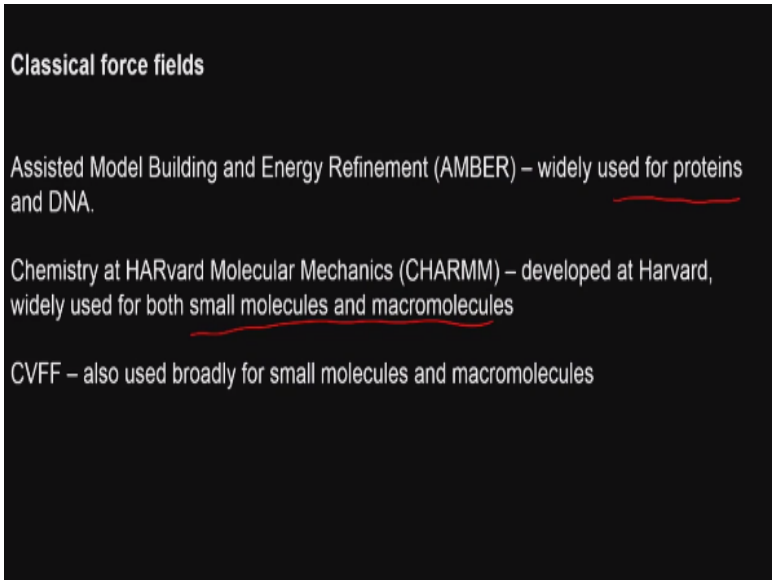
(Refer Slide Time: 09:44)



So there are many force fields like I said MM2 this is developed by Norman Allinger mainly for hydrocarbons small organic molecules has a large set of parameters like I showed you in that table large set of parameter that is continuously refined and updated for many different classes of organic compounds, so MM2, MM3, MM3+, MM4 and so on. So this is a very good force field if you are looking at drugs organic molecules, okay, without any metals or something.

Then of course CFF okay, this is developed by these people it is a general method for unifying studies of energies, structures, vibrations of general molecules and molecular crystals. The program was developed by these 2, it is based on the Cartesian representation of the all the atoms, okay, and then based on the CFF many other force field of these family consistent came into, I will show some of them.

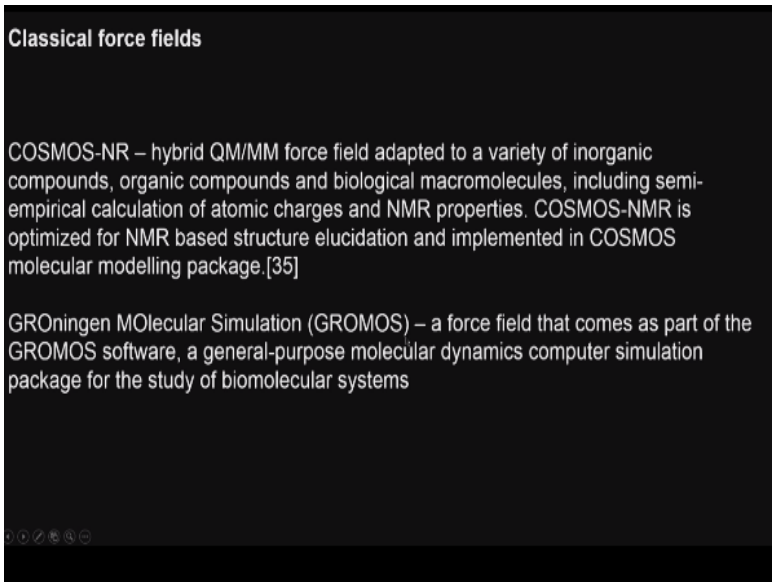
(Refer Slide Time: 09:49)



So more of them assisted model building and energy refinement. This is called AMBER force field; this is widely used for proteins so you see if you have a lot of protein we can use this. Then comes CHARMM, CHARMM is another very popular force field used for small molecules and macromolecules so generally MM type of force fields are used for organic drugs, AMBER is used for proteins.

CHARMM is used for small molecules and macromolecules this is the rule of thumb, so you can use MM type of force fields, you can use AMBER or CHARMM okay depending upon whether small molecule drugs, macromolecules or whether it is a protein. CVF is broadly used for small molecules and macromolecules, there is some software used at CVFF, there is a consistent valance force field.

(Refer Slide Time: 10:47)

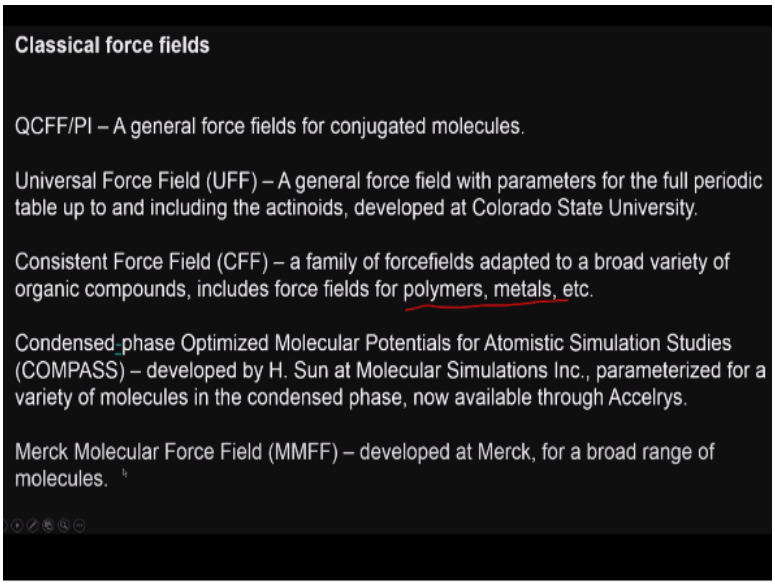


This is COSMOS NR, it is a hybrid of quantum mechanics and molecular mechanics so it combines little bit, adds little bit of quantum mechanics adapted to a variety of inorganic compounds okay, so if you are looking at inorganic, organic compounds, biological macromolecules including semi-empirical calculations atomic charges NMR, so COSMOS NMR is optimized for NMR based structural elucidation, okay remember that.

In drug discovery we do not spend much time on that actually. This is again GROMOS, it is a free software, it can be used for lot of dynamic simulations, this is called GROninger Molecular Simulation a force field that comes as part of this software, it is a free software for dynamic study, it is molecular dynamics, computer simulation package for the study of biomolecular systems.

Okay, so if you are using that software we will be using this force field, okay, more force fields;

(Refer Slide Time: 11:47)



OPLS, optimized potential for liquid simulation, so there are many OPLS-AA, OPLS-UA, OPLS-2001 developed by Jorgensen at the Yale University Department of Chemistry, it is very good for liquid simulation, okay. Then ECEPP, first force field for polypeptide molecules developed by Momany, Scheraga and colleagues, then QCFF/PI, this is a general force field for conjugated molecules.

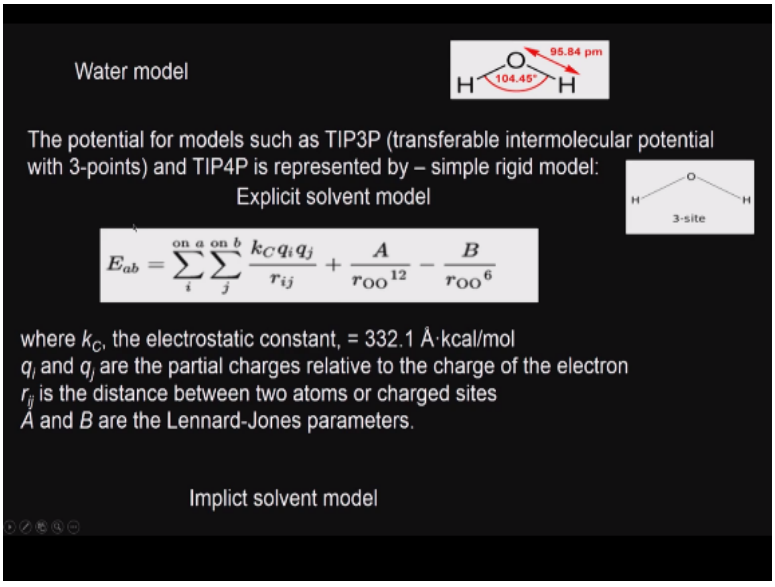
UFF like I said universal force field they call it general purpose force field parameters for the full periodic table including the actinoides developed at Colorado State University, so they try

to have parameters for all the atoms in the periodic table, okay then the consistent force field so these are family of force fields adapted to broad variety of organic compounds, force fields for polymers.

Okay so if you are interested in polymers, metals, then maybe you can go for this type of force field, okay. Then we also have condensed phase optimized molecular potentials for atomistic simulation studies, COMPASS, this is developed by these particular Sun computers at molecular simulation incorporated for a variety of molecules in the condensed phase now available through Accelrys.

Accelrys is a software, commercial software so if you are having that software you will be having this particular also, okay, Merck Molecular Force Field developed by Merck company MMFF for a broad range of molecules, as you can see some of these force fields are specifically developed for their internal use which might not be available for general public whereas things like MM2, MM3, CHARMM, AMBER, CFF you may get it as a free ware.

(Refer Slide Time: 13:43)



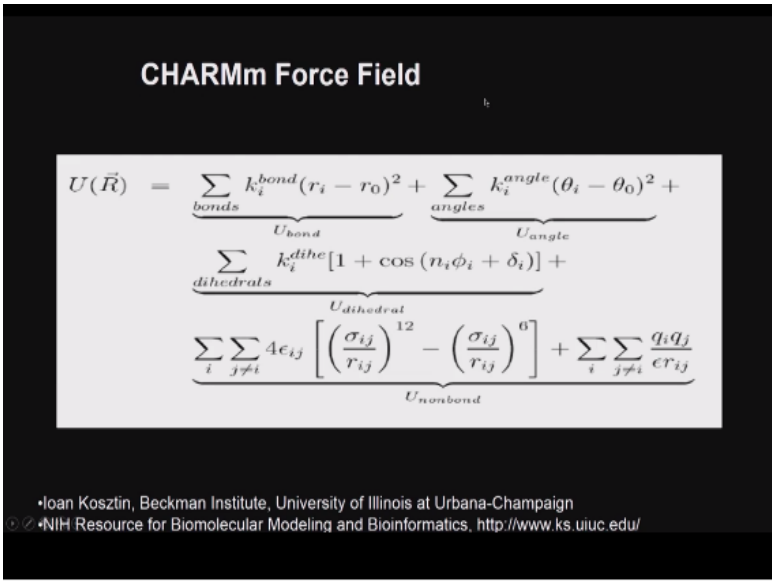
Okay, water, water plays a very important role because when you have proteins water will be found, water molecules will be found, so modeling water also has become a big issue nowadays, so should I have implicit solvent that means I have a change in the dielectric constant depending up on water or air or should I have explicit, so water as you know it is a very interesting system, it has got 3 sides okay, so these and these can form the hydrogen bond donor.

This can be hydrogen bond acceptors okay, the distance is given and angle is given here so TIP3P transferable intermolecular potential with 3 points like 3 points this this this, and TIP4P is represented by simple rigid molecule, so the model looks like this, okay, there is a 12 here, 6 here, then there is a constant kc, which is the electrostatic constant given by this.

Qi qj are the partial charges related to the charge of the electron, rij is the distance between 2 atoms or charged sides. A and B are the Lennard-Jones potential term, so imagine if you have lot of water present, the calculations increase because we need to add some extra terms. We need to consider water because water plays a very important role in hydrogen bond formation, donor as well as acceptor, ligand protein interactions, changes in the conformation all these happens because of the water.

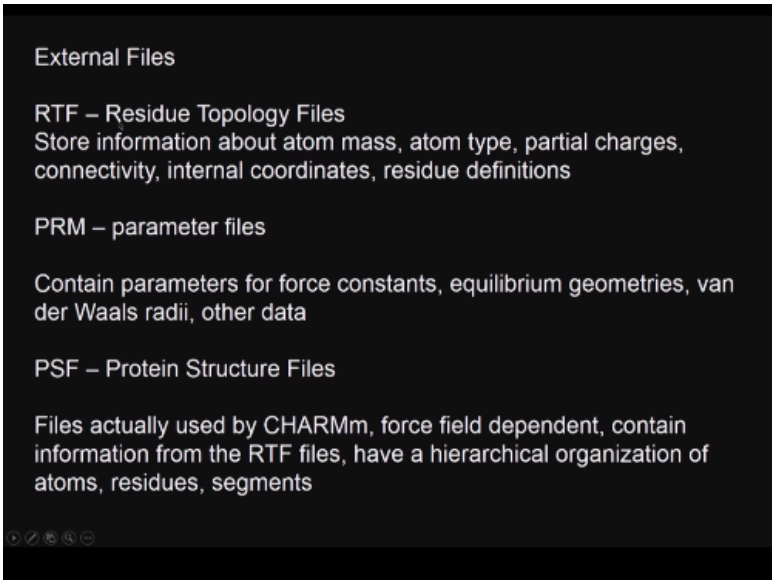
Although the energy is as I showed you long time back is much much < 10 kilo calories, but it plays a very important role in the conformation, the ligand takes with respect to the active site when it gets bound to the protein, okay, so water is a very important parameter we need to consider, okay, let us look at some functional forms of these force field, CHARMM.

(Refer Slide Time: 15:50)



As I said CHARMM is widely used force field and as you can see this term is the bond stretching I showed you long time back they have a term for angle bending, they have a term for dihedral okay, then they have a term for electrostatic and Lennard Jones potential, okay, so this is what CHARMM used very whatever I have been talking about, okay, they have very clear simple looking force field.

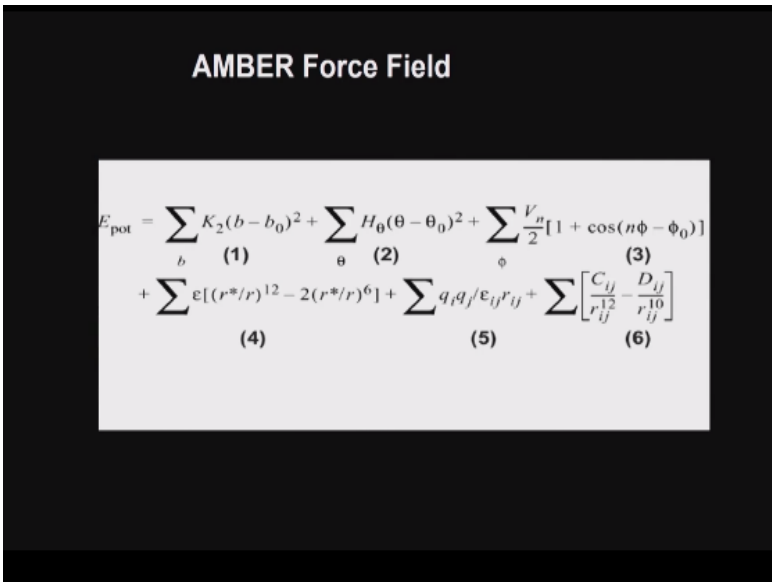
(Refer Slide Time: 16:20)



There are files, external files RTF, residue topology files, this stores information about atom mass, atom type, partial charges connectivity, internal coordinates, residue definition, then we have the parameter files contains parameters for force constants, all these parameter, okay, equilibrium geometries, van der Waals radii, other data, so equilibrium geometry is the geometry it takes and it is the equilibrium state.

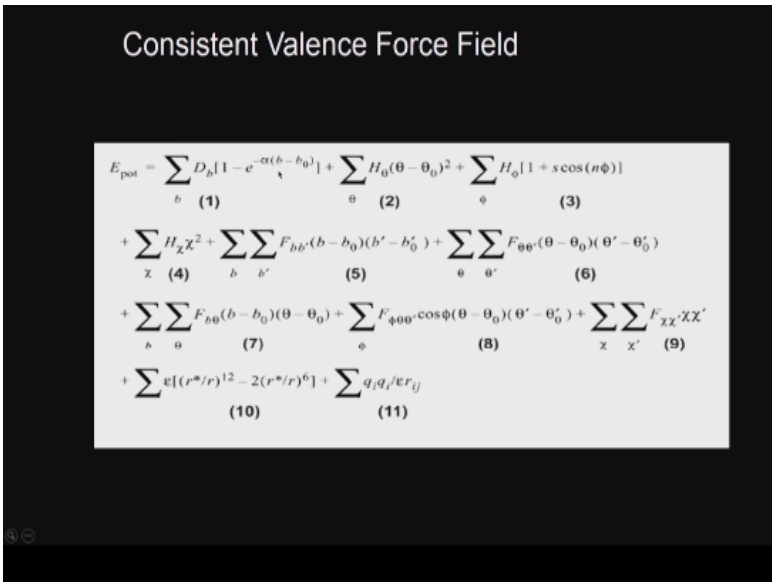
Then there is another file called protein structure file, files actually used by CHARMM force field dependent contains information from the RTF files, okay, have a hierarchical organization of atoms, residues, make segments so these are the files that are necessary if you are running a CHARMM force field okay.

(Refer Slide Time: 17:15)



Then comes AMBER, look at this AMBER, AMBER also looks very much similar, you have the term for bond stretching, bond bending, torsion okay, Lennard Jones potential, the electrostatic and of course you also have the hydrogen bond term also that is one extra.

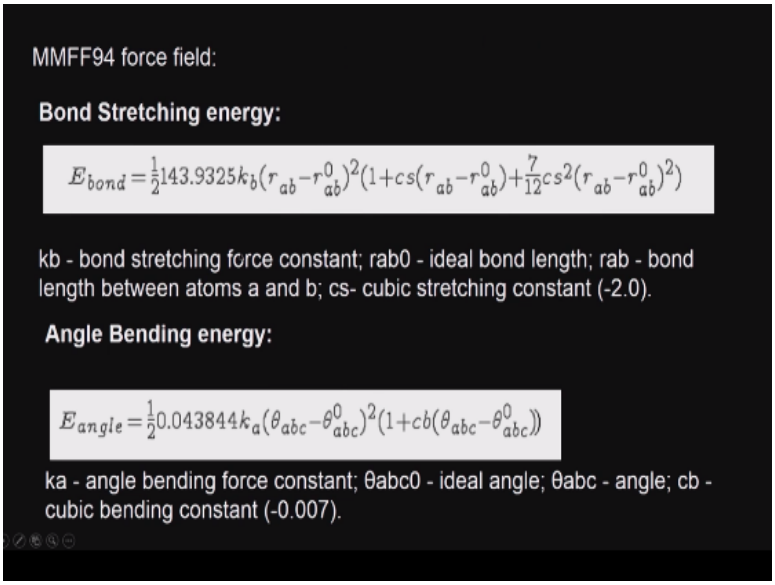
(Refer Slide Time: 17:35)



Then CVFF consistent valence okay, as you can see the bond stretching has become different, Morse term not this type of term okay, they use the Morse term then they use the same equation for the angle bending, torsion, okay, now you see there are lot of cross terms there okay, stretch-stretch, angle-angle, stretch-angle like that you know, lot of cross terms, and then of course this is your electrostatic and this is Lennard Jones, okay.

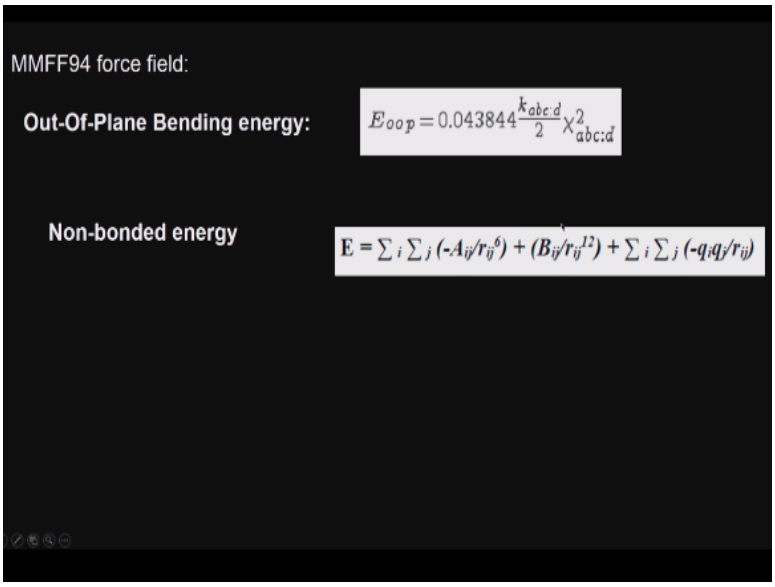
So CVFF uses many cross terms as you can see here, okay, so obviously we need extra parameters, these are extra parameters which require, if you are going to use cross terms okay, and is called CVFF force field.

(Refer Slide Time: 18:33)



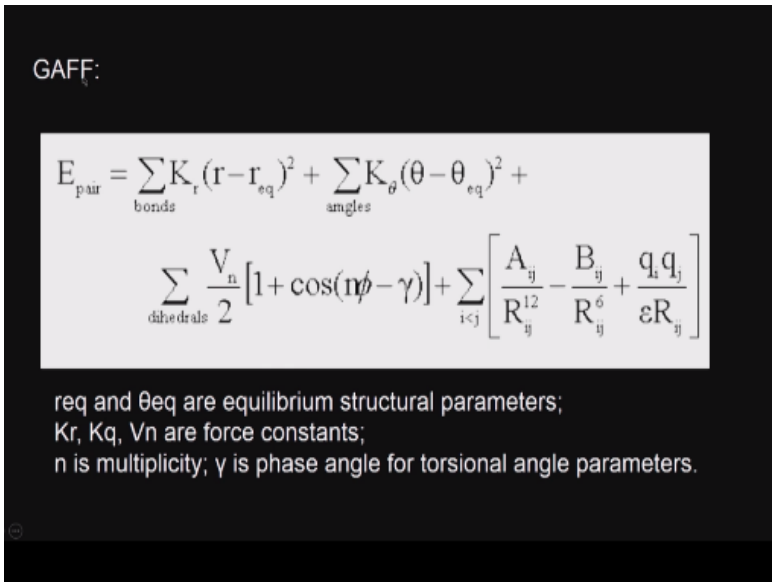
Another force field MMFF94, look at the equation for bond stretching, okay, so they have completely different equation, they consider square terms as well as linear terms okay, as you can see here and you can see these are constant then for the angle bending also you can see square term and a liner term okay.

(Refer Slide Time: 19:00)



Out of plane bending energy, I talked about if you have 4 atoms, 3 maybe in one plane, the fourth one maybe out of the plane, so that is also included here, these are non-bonded interactions okay, the electrostatic and the Lennard Jones potential. So MMFF uses okay not only the quadratic term, but also the linear terms as you can see here okay.

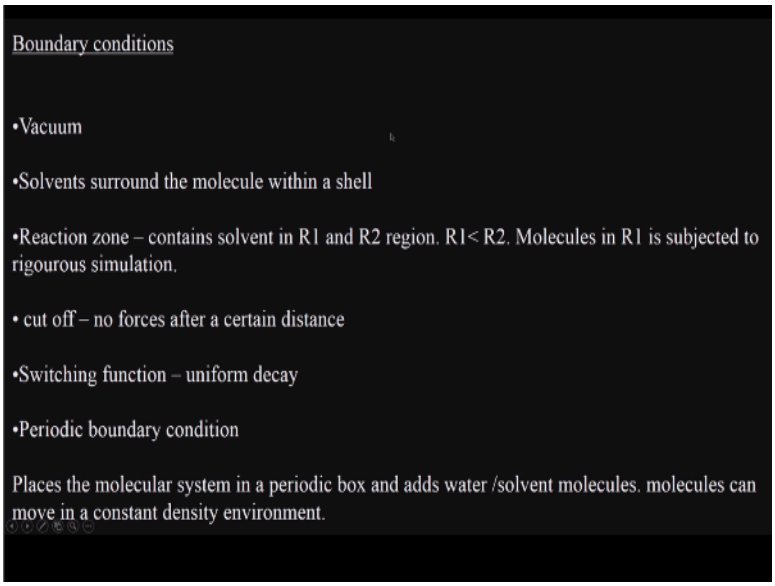
(Refer Slide Time: 19:31)



This another force field called GAFF okay, if you look at the GAFF, so we have the stretching, using the normal square term, bending and then torsion and then they have the standard term for electrostatic and Lennard Jones, so different types of force field terms so that is why in the beginning I mentioned that never mix and match force field, so always use the same force field.

So as you can see the functional forms can be very different between force fields and of course the corresponding parameters also will be very different between each force fields. So always use the same force fields if you are studying, a set of system and if you are trying to compare a set of system okay.

(Refer Slide Time: 20:16)



So we know how to use the force field which one to select and what is the difference between these force fields. Then comes the boundary conditions, what is this boundary condition? So if you have a molecule I want to study the molecule so the molecule maybe surrounded by some solvent or water or it could be in vacuum.

For example, if you are studying the energy or conformation of a molecule which is very dilute form or if it is in a gas form we can consider it as vacuum okay, so there is no solvent interaction, there is another molecule of that same type is also not there because if it is in gas phase it may be very far apart, okay, so it has become easy, we do not have to consider molecule-molecule non-bonded interaction.

Or if you have a solvent then there will be solvent molecule interaction okay, so if I have to say chloroform there will be say 100s of chloroforms surrounding my drug then there could be lot of interactions between chloroform molecule and the drug, non-bonded interaction, electrostatic, van der Waals and so on actually or hydrogen bonds. So when you consider vacuum calculations are easy.

We just have that molecule and study it when you have solvent you may surround it with water or solvent or chloroform or methanol so you are going to have lot of non-bonded interaction between the molecule and the solvent, so how many solvents do I consider layer, how many layers, so shell so we can consider between R1 and R2 okay or we can have a cut off that means no forces after certain distance.

Because if you consider solvent up to certain 3 layers or 4 layers then after the fourth layer it is like vacuum right. So we need to balance and of course if you have too many layers of solvent then the calculations also become too many, you may have almost factorial extra calculations to be done to determine all the non-bonded interaction so to be more realistic you may have many layers of solvent like your molecule.

And there will be solvent-solvent, solvent-solvent or the calculations become too many or do I make it little bit approximate and assume only few layers of solvent so that is the question we need to answer. There are different approaches called uniform decay so the interaction of this is much more than the interaction of this, interaction of this so there is a uniformly decreasing level of interactions that is called the uniform decay.

Periodic boundary conditions, we will talk about this more in detail later, so periodic boundary condition places the molecular system in a periodic box and adds water or solvent molecules. Molecules can move in a constant density environment, basically it maintains the density of the entire system of the molecule with the solvent so different approaches can be followed.

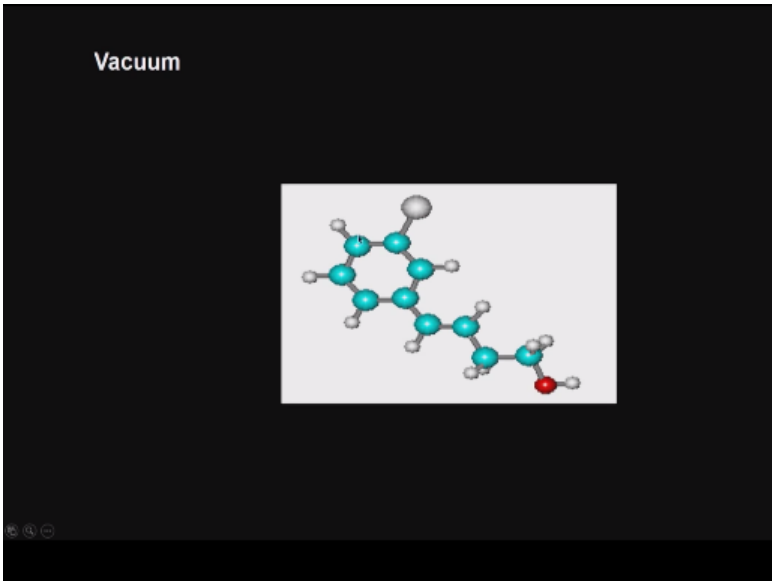
I can think about vacuum that means I have only my molecule nothing else, so there is no interaction with anything else, solvent I could consider solvent surrounding my molecule so there could be non-bonded interaction between the solvents and the main compound or hydrogen bonds okay, so how many layers of it do I consider that becomes a challenge actually.

Do I just consider 1 layer or 2 layers or 3 layers or 4 layers so the calculations become more intensive if I consider more and more and more layers okay, another approach is uniform decay of the interaction of the solvent with the main compound. Then another approach is this periodic box, this is a very important concept which is generally used when we talk about proteins.

So what it does is it places the molecular system inside a box containing water or solvent molecules okay so density is maintained constant, so if a solvent goes out of the box then another solvent is assumed to enter from the opposite direction okay, that is what is called periodic box. We will talk about this little bit as we go along now. So different types of boundary conditions need to be considered.

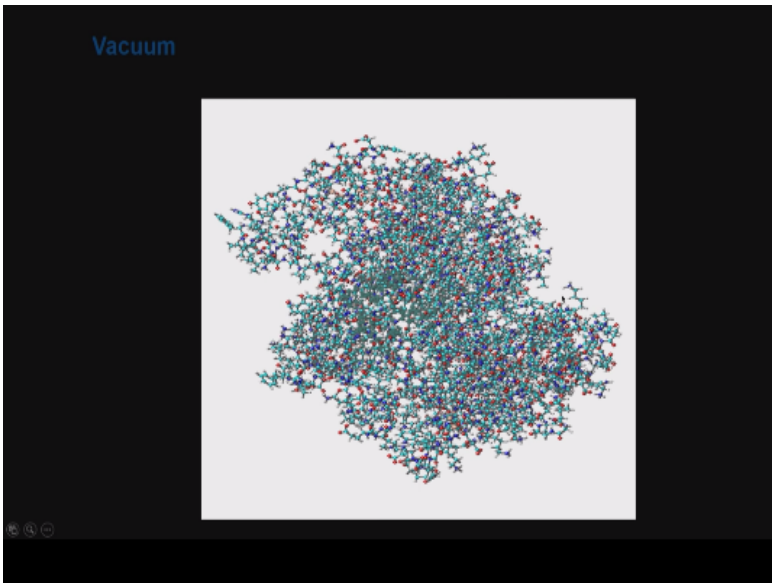
As you know proteins need to be in water environment so that it can assume proper conformation okay for example if protein is surrounded by water the hydrophilic portions may come out, the nonpolar hydrophobic regions may go inside, so the conformations can change dramatically depending up on the solvent or the boundary conditions, okay, so for example vacuum like I said.

(Refer Slide Time: 25:22)



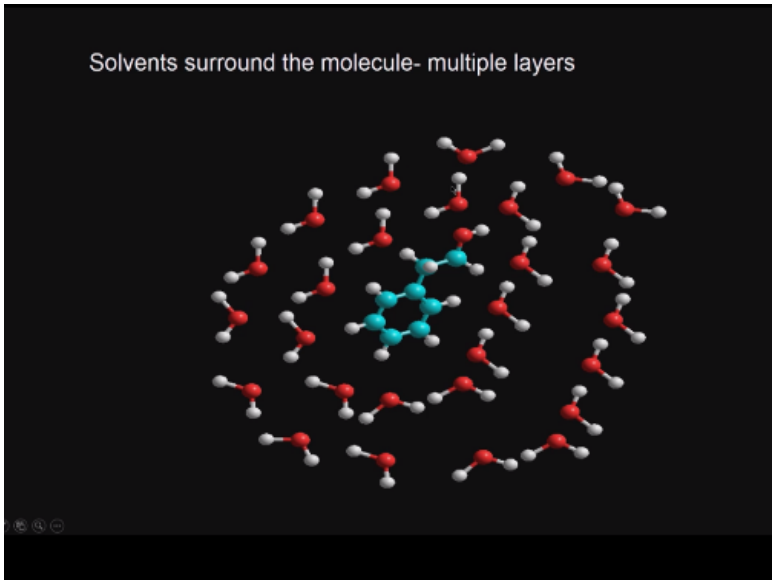
I may have just the molecule and I may look at the energy of this molecule and what conformation this molecule takes place, so I can study this, okay, I can calculate shape, size or heat of formation of this molecule, there is nothing else surrounding this all alone so it is like a vacuum or gas phase.

(Refer Slide Time: 25:41)



Or I can have molecule say protein like and there is nothing surrounding this although it is not realistic because proteins generally are surrounded by water because they are always found in aqueous medium.

(Refer Slide Time: 25:59)

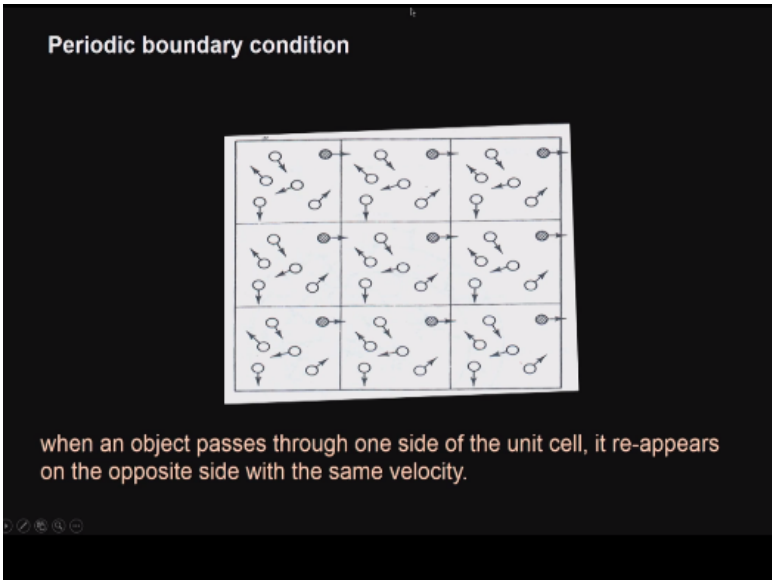


So I can think of a solvent layer surrounding this main molecule so obviously there is going to be lot of hydrogen bond accepting and donor happening okay between all these water molecules and solvent so the total energy of the system can be different the energy of this molecule itself can change, the molecule can take new conformations because of that. I can consider 2 layers of solvent so as you can see the amount of calculations increase.

But it is more realistic or I can consider 3 layers of solvent okay, but then some interesting thing is the solvent at the outer most layer has interaction with the solvent in the inner layer, but outside that is the vacuum, so these solvent face a very non-realistic forces, okay, so there are many approaches to handle that. We will not go into that especially the solvent in the outermost layer has this problem okay.

Whereas the main compound has no problem because you have surrounded one layer, one layer, one layer and so on okay.

(Refer Slide Time: 27:10)

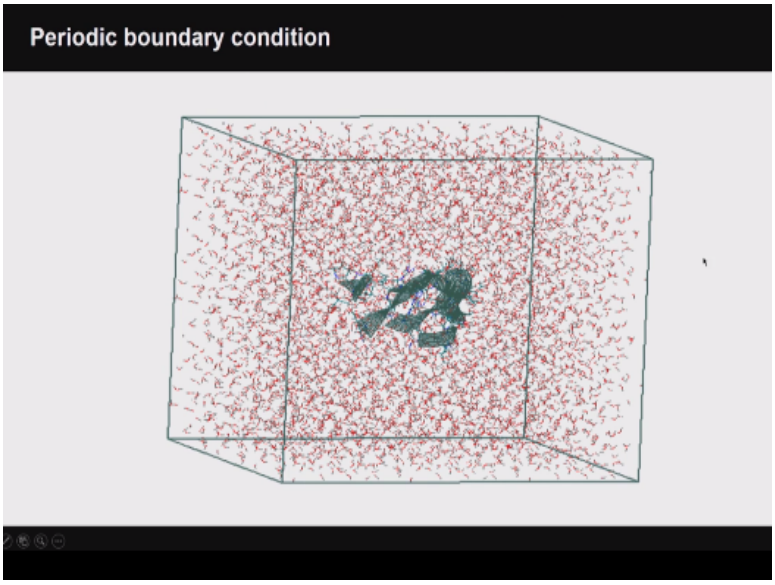


Periodic boundary condition, let us look at this, what is this? When an object so we have periodic box okay, and we keep the solvents so when the solvent, one solvent goes out from the box we put in another solvent from the opposite direction so the number of solvent molecule inside the box is constant so the density is constant and if it is a square then we place similar looking boxes in all sides, okay.

So we do the calculations for all. If it is a cube then obviously you have 4+2=6 sides of it so everywhere you assume another cube placed and you do the same calculation, so the density of the inner most cube remains constant because if one solvent goes out during the calculation we assume another entering from the opposite directions. The number of solvent molecules are constant.

So periodic boundary condition is generally preferred which is much more accurate for considering the surroundings, okay, for considering the boundary condition okay for example.

(Refer Slide Time: 28:26)

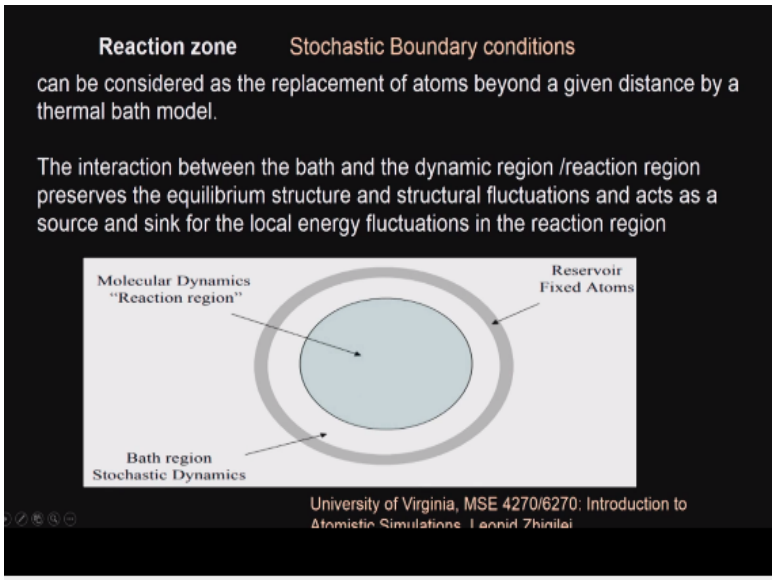


This is the protein, this is the periodic box, we have assumed it here okay, a large cube like, there are lot of water molecules here and so we assume similar boxes on all the 6 sides of this cube so we can perform molecular dynamic simulations we can perform conformational changes so if one water molecule goes out from this box to the box outside it is assumed another water molecule will enter.

So the number of water molecules inside will be always constant. So the density is constant okay and why it is periodic because you have assumed that there are 6 more boxes of similar type on all the 6 faces of this cube. So periodic boundary condition is generally preferred like I said calculations are intensive as you can see here, if I take a protein I may have about 2000 or 3000 water molecules.

So we are calculating so many non-bonded interactions, hydrogen bonds with the main molecule so calculations are quite a lot, okay, I also mentioned reaction zone that is another approach by which we can incorporate boundary conditions.

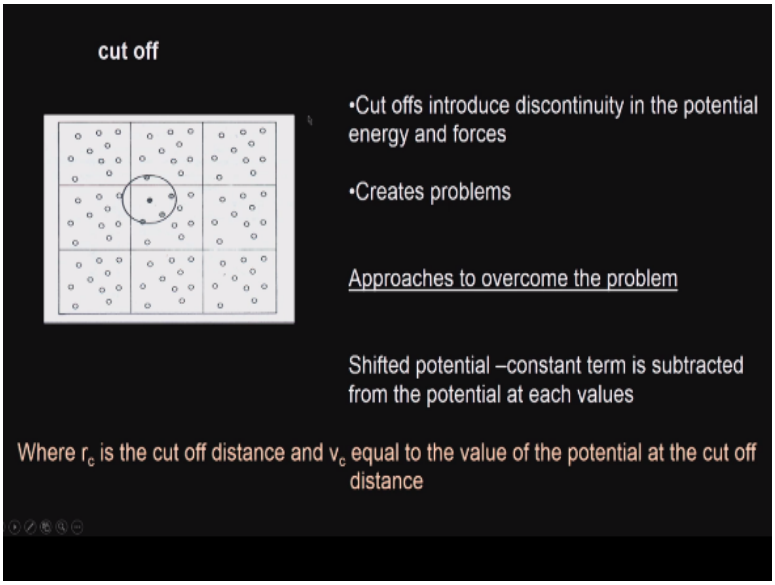
(Refer Slide Time: 29:45)



Reaction zone, this can be considered as a replacement of atoms beyond a given distance by a thermal bath model, so we have the reaction region, we have the stochastic dynamics here, you have fixed amount of atoms here okay, the interaction between the bath and the dynamic region, reaction region preserves the equilibrium structure okay, and then this reservoir okay, maintains the energies these solvent molecules face actually okay.

That is called the reaction zone, it is a stochastic boundary condition. Another approach is called the cut off okay.

(Refer Slide Time: 30:25)



So cut off introduce discontinuity in the potential energy and forces, okay, so beyond a certain distance we do not assume any solvent present so up to that so solvents here interact with the main compound and this portion is assumed as a vacuum so obviously there is

discontinuity okay, beyond this so it can create problems so how do they do that, there are several approaches.

We will not go into that okay, shifted potential that is constant term is subtracted from the potential at each value, so different types of approaches are used so that this type of large discontinuity does not happen, another approach which is said is uniformly decaying function that means the solvents surround immediately has certain energy of interaction, solvents beyond that, you have a uniform decay and as you go out and out and out the effect of the solvents on the outer layer is much less.

So you get a uniform decay rather than sudden cut off, okay, that is sudden cut off which you say as soon as the cut off here solvents here do not contribute anything to the main compound, okay, so we shall continue further on this molecular mechanics and force field in the next class as well. Thank you very much for your time.