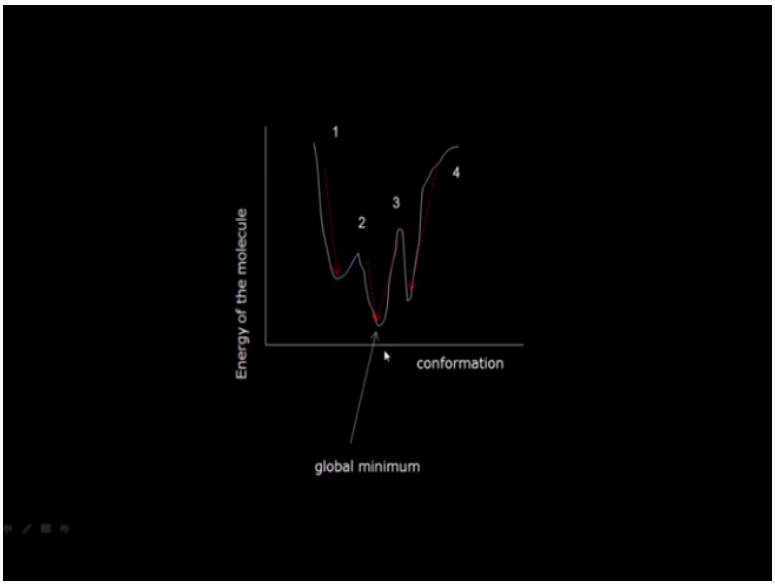
Computer Aided Drug Design
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture - 22
Conformational search/MD

Hello everyone, welcome to the course on computer aided drug design. We have been talking about force fields, molecular mechanics then we looked at different types of force fields, then we looked at conformational search and then we also looked at some numerical methods for solving differential equations. So, we will continue little bit on this conformational search. Because, that is a big challenge, how do you know that the molecule conformation is equivalent to your minimum energy and that is what we have been looking at in the past few classes okay.

(Refer Slide Time: 00:48)

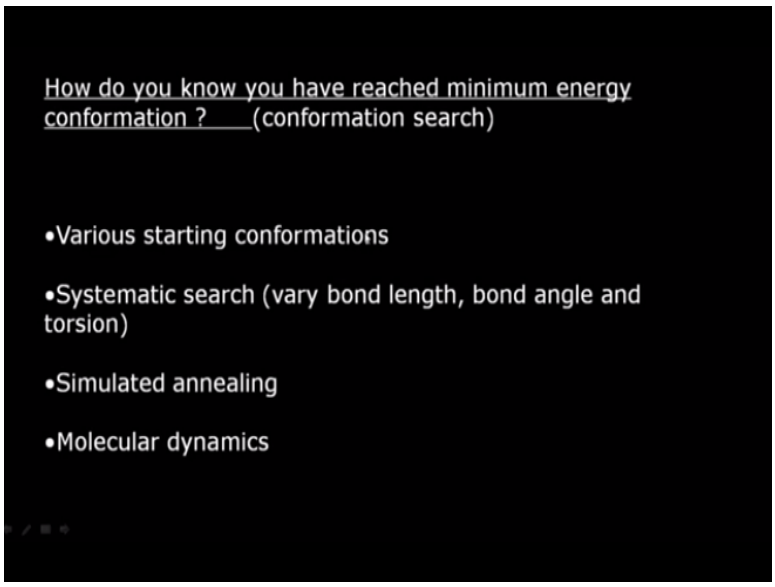


So, basically you may start at different places and you may end up at some minima. But, this is a local minima. Whereas, this is the global minima. So, depending upon where you start you may end up but, your goal is to reach this global minima. So, conformational search is a big challenge for small molecules, it is okay but for large molecules it becomes a big problem. And as you know, the biological activity sometimes depends upon the conformation.

Whereas when it goes and binds into an active site, it may have a different conformation depending upon the flexibility of the bonds and so on. So, there are different types of

conformational algorithms. So, how do you know you have reached a minimum energy okay. I mentioned that, you can look at the change in energy or you can look at the delta e/delta change in the conformation. But that could leading to local minimum rather than global minimum.

(Refer Slide Time: 01:57)



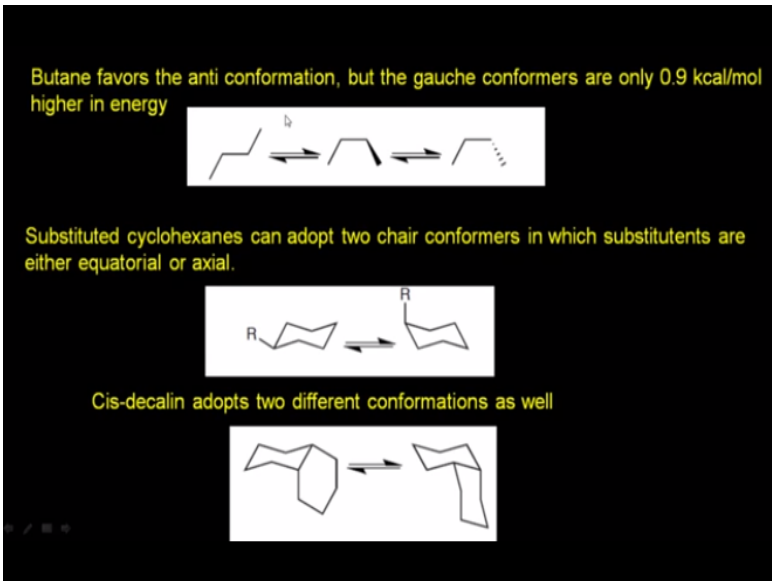
But, some of the approaches that are followed to find out whether you have reached global is to start with various starting conformations. For example, suppose I start with 1, I may end up with a minimum, which is not global but local. Then I start with 2, I will go down much lower than this, then I may start with 3, which may go here, I may start with 4, which may go here.

So, by starting with different conformations and then keep minimizing and then look at the minimum of this minimum and that we call it the global minimum okay. That is starting with various starting conformation. You can also have systematic search, that means you change bond length, bond angle, torsion very systematically keep on calculating the energies but that is going to be very time consuming because, when you have say 10 bonds and 20 angles and 30 torsions imagine, it will be in almost like raise to the power.

Another approach is called simulated annealing that means you heat the molecule to higher temperature because of the energy, the molecule as it takes lot of conformations and hopefully when you minimize those conformations it will lead to a global minimum. And then another approach is molecular dynamics by which we can also get conformation. We will look at some of them little bit in more detail.

So, the main challenge here is to reach the global minimum not end up with local minimum, that is the main challenge okay.

(Refer Slide Time: 03:29)

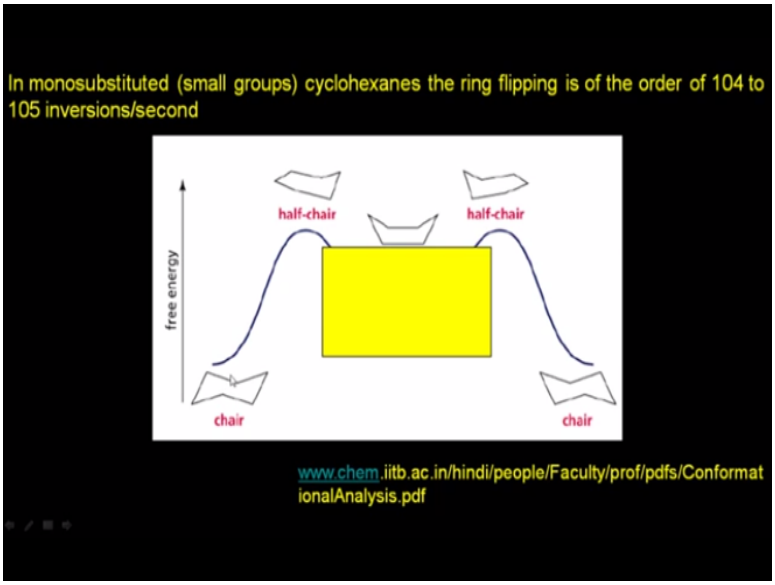


For example, if you look at butane. Butane, as you know is c4 and it can take this type of conformation or this type of conformation. That means, here the bond is sticking out of the paper, whereas here, the bond is going inside the paper. So, but then, there is only small change in the energy 0.9 kilo calories okay. So, you cannot say all the 100% will be in the minimum energy conformation.

We will have ratio because the change is only little. Look, substitutes like cyclohexanes, this is substituted cyclohexane okay and the cyclohexane as you know can take chair conform or boat and all that right. Substituent's are either equatorial or they are axial okay. So, they can like this or like this. Look at Cis-decalin, decalin as name goes 10 okay. So, we have 1,2,3,4,5,6,7,8,9,10 okay.

So, you have one close ring another close ring. So, it can take 2 different conformations okay like this and like this.

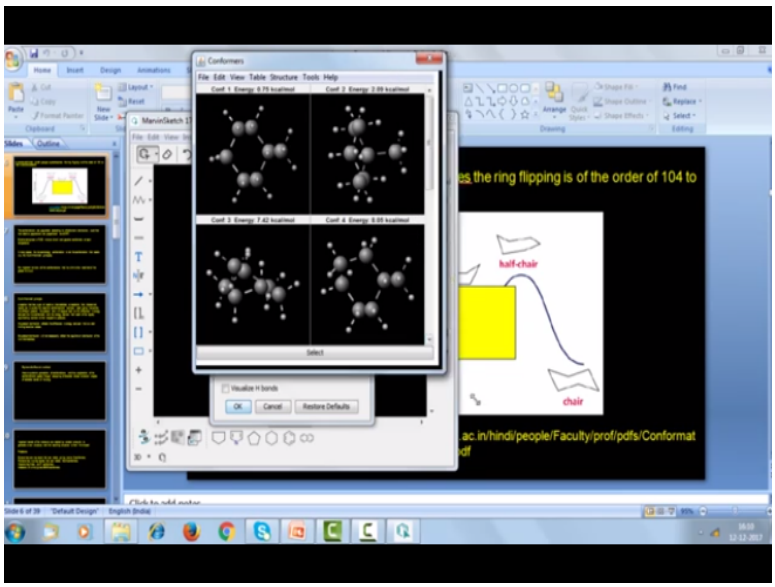
(Refer Slide Time: 04:43)



So, if you look at cyclohexanes we all have studied it can take chair form or boat form and so on right okay. If you look at monosubstituted cyclohexanes, the ring flipping is of the order of 104 to 105 inversions per second. So, it can change quickly and this is the free energy on the y axis so this can be half chair, this is full chair, again half chair, this is again chair. So, it can be flipping this way and that way.

This is taken from this particular reference. So, as I mentioned before, the change in the energy is only 0.9 kilo cals. So, it is not a big deal actually okay. So, it is not a big deal for us.

(Refer Slide Time: 05:30)



So, as I had mentioned before, we can look at the Marvin sketch and as we can also get the for example, this is cyclohexane and this is substitute at cyclohexane okay. So, we can say structure, clean in 3D so, we have a 3D version of this and as you know okay. So, this is for

this and we can look at it like this okay. So, this is not the minimum energy conformation so we can look at Marvin sketch as you know here we can view it in different forms I had mentioned it before.

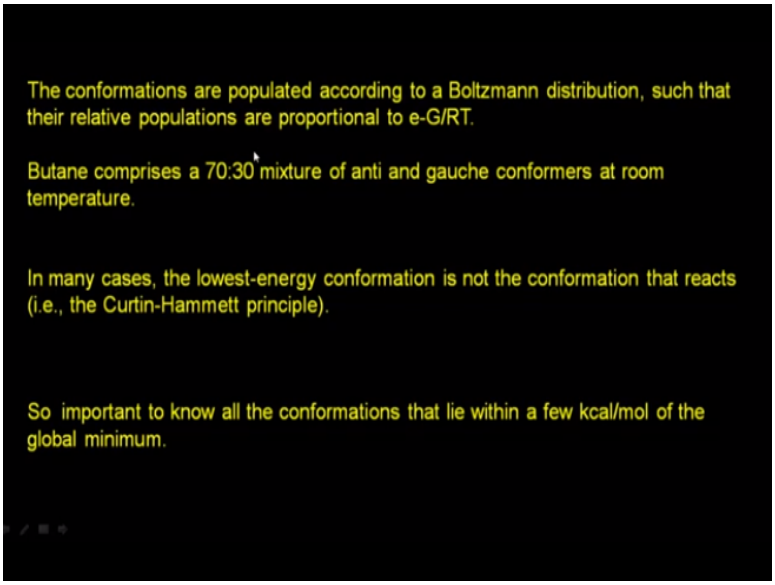
So, it can be a wireframe type of form okay small and stick okay. Now, we can look at some calculations, we can look at conformations, different conformers we can say okay. So, I want to look at 10 conformers here. Look at this so, it is giving you different conformations of this substituted cyclohexane so, this is 9.75 so this is very, very different okay. So, forget about this, this is 7.42, this is substituted, this is 8.05 okay look at this conformation okay.

As you can see this is nice looking chair. So, it gives you different conformations of the molecule. Look at this, this is 27. So, definitely very high energy, whereas look at this conformation 2.09 okay. This is 7.42 conformation, this is 0.75 still lower than this okay. So, you get a beautiful picture here on this. So, lot of conformers with different energy okay. So, if you want we can calculate with the MMFF energy optimization limit is normal so we can give a strict optimization also.

So, by doing that so we can see this particular conformation is quite low, energy wise energetically okay. This gives you the lowest energy which is 0.7 kilo cals and this particular gives you quite a high energy 8.3 kilo cals okay. So, as you can see that is torsion whereas the distortion here is much lower. Here the distortion is much higher okay whereas here you can see a nicely switch sane okay.

The carbon sitting here, that is why the energy is very low 0.7 kilo cals. So, you can see from this software also, we can get different conformations of the molecule here. So, we can go into conformation, conformers so we see 13 kilo cals, 13.5, 13 so some of them are good 14 so, we get different conformations. But, most of them are almost same very little differences in these conformations because the software tries to minimize the energy okay. So, there is a very important.

(Refer Slide Time: 12:32)

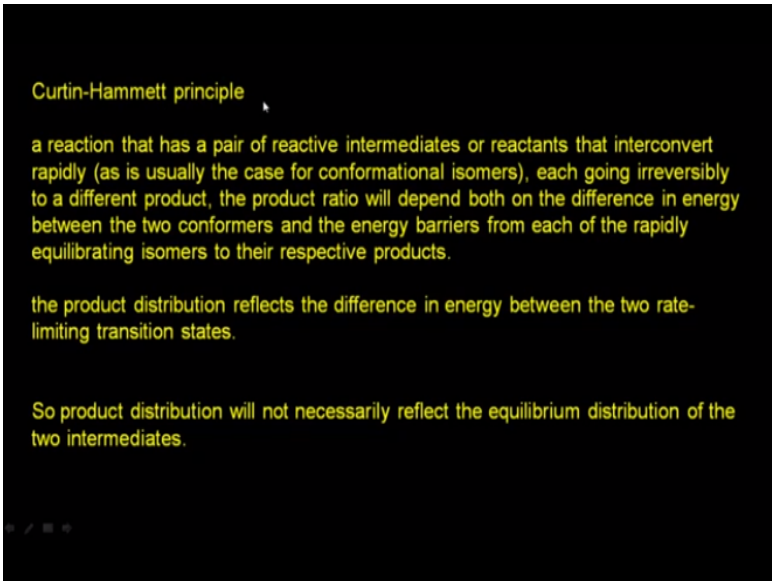


The conformations are populated according to Boltzmann distribution, such that their relative populations are proportional to e-G/RT okay. So, it is not that you will always find minimum energy conformation. For example, butane, you will find both anti and gauche in the ratio of 70:30 okay. Because, their distribution will be based on e-G/RT. In many cases, the lowest energy conformation is not the conformation that reacts okay.

So, the biological reaction need not be the lowest energy. And there is a rule called Curtin-Hammett principle, we will look at that. So, important to know all the conformation that lie within a few kilocalories. Like I showed you in some of these examples so, just do not go and stop with minimum energy. We will look at conformations with + or - 2 kilo cals look at all those conformations because as you know, the conformations are populated according to this particular distribution okay.

So, the energy is higher we will have less of it but, still you will find some if energy is lower you will find more of it. Like for example butane, you will not find only the minimum energy but you will find the other one also. But, the ratios will be different. Now, what is this Curtin-Hammett principle?

(Refer Slide Time: 13:54)

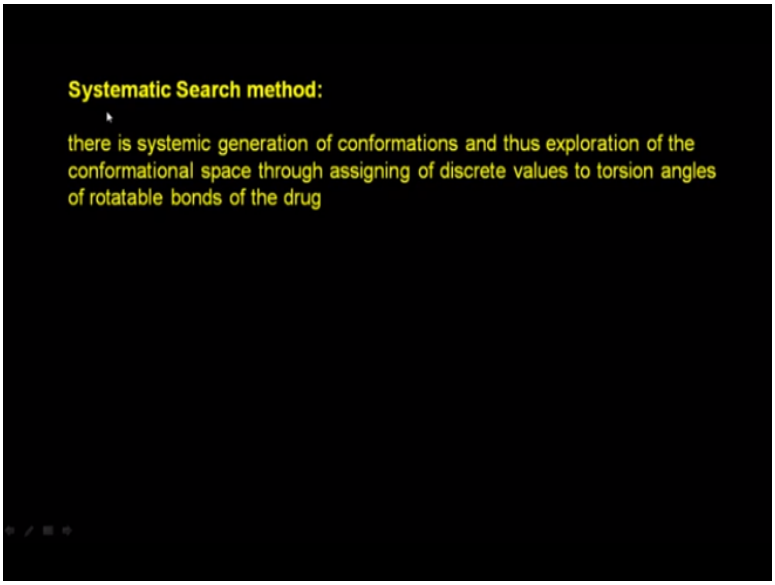


What is says is, the reaction that has a pair of reactive intermediates or reactants that interconvert rapidly. So, if you have say A, B 2 different conformations, they interconvert between those 2 conformations very fast because the energies are that is conformational isomers. The energies are not very differently high. Each going irreversibly to a different product. So, they quickly shift from one conformation to another but then, they go into the product.

So, the product ratio will depend both on the difference in energy between the 2 conformers and the energy barriers from each of the rapidly equilibrating isomers okay. So, if I have a particular ratio, it is not that I will get the same ratio as the product but, it will also depend on the difference in energy between the 2 transition states. So, the product distribution will not necessarily reflect the equilibrium distribution of the 2 intermediates remember that.

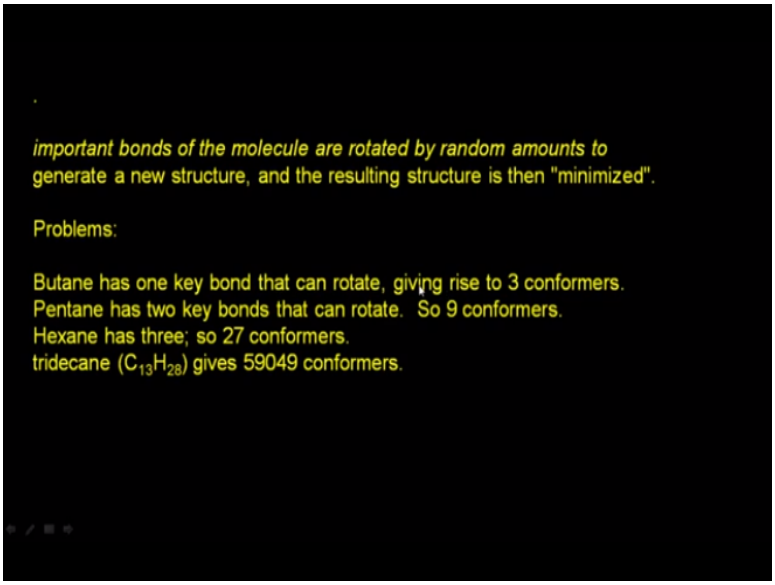
So, 2 conformers interconvert rapidly, there is very small difference in that energy and then they go into product, that which is irreversible. Because, they can interconvert rapidly in equilibrium, the product distribution will be based on this difference in energy between the 2 rate limiting transition states okay.

(Refer Slide Time: 15:20)



So, one of the approaches is, like I said systematic search method. There is systematic generation of conformations. So, there by explore the conformation space through assigning discrete values to torsion angles rotatable bonds and all that.

(Refer Slide Time: 15:40)



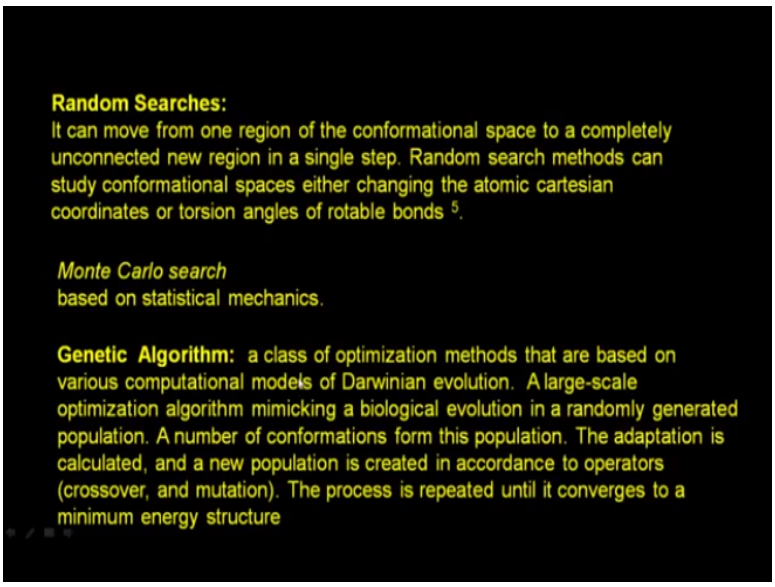
It is not so easy important bonds of the molecule are rotated by random amounts to generate a new structure then the resulting structure is minimized. For example, butane has only one key bond that can rotate so, it can give 3 conformers. Pentane has 2 key bonds, pentane means C5 okay 1,2,3,4,5 pentane. So, pentane has 2 key bonds okay. So, this bond and this bond pentane has 2 key bonds. Whereas, if you look at butane has only one key bond okay.

Butane for example, this is the only one key bond for butane which can rotate. Pentane has this bond and this bond, 2 key bonds which can rotate. So, in the case of pentane you can

have 9 conformers. Hexane has 3 key bonds okay. Let us look at hexane 3,4,5,6 so, hexane has 3 key bonds 1,2,3 so we will have 27 conformers. Tridecane that means C13 decane is 10 tridecane C13 will give you 59049 conformers.

So, you see it is not so easy, there are so many rotatable bonds as the carbon number increases so, you may have to have look at so many conformations okay. As you can see from butane to pentane to hexane to tridecane, number of conformers increase rapidly.

(Refer Slide Time: 17:38)



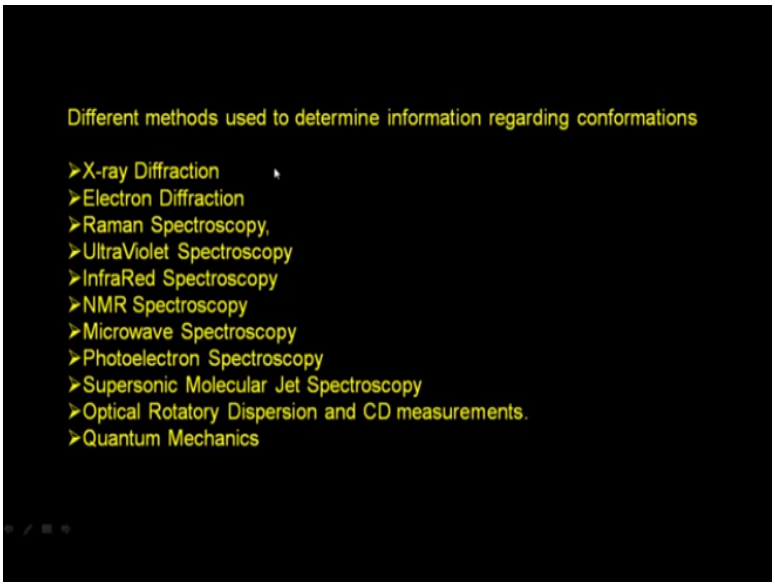
Then you have random search, that is another approach. It can move from 1 region of the space to completely unconnected new region in a single step. So, you randomly change things. You may change the Cartesian coordinates okay. Cartesian coordinates means x, y, z or torsion angles of rotatable bonds. You go another approach that is called Monte Carlo search, this is based on statistical mechanics.

Then we have the genetic algorithm, these are called a class GA. A class of optimization method that are based on Darwinian evolution. So, a large scale optimization algorithm mimicking a biological evolution in a randomly generated population. So, you look at number of conformations from this population okay. Then you look at those population, the adaption is calculated and a new population is created in accordance to operators crossover mutation okay.

So, we keep on. You look at a set of conformations, you take those which look good, then eliminate the ones which does not look good. So, that is almost like evolution and then

proceed okay. That is what is called genetic. There are various variations to genetic algorithm also okay.

(Refer Slide Time: 18:58)

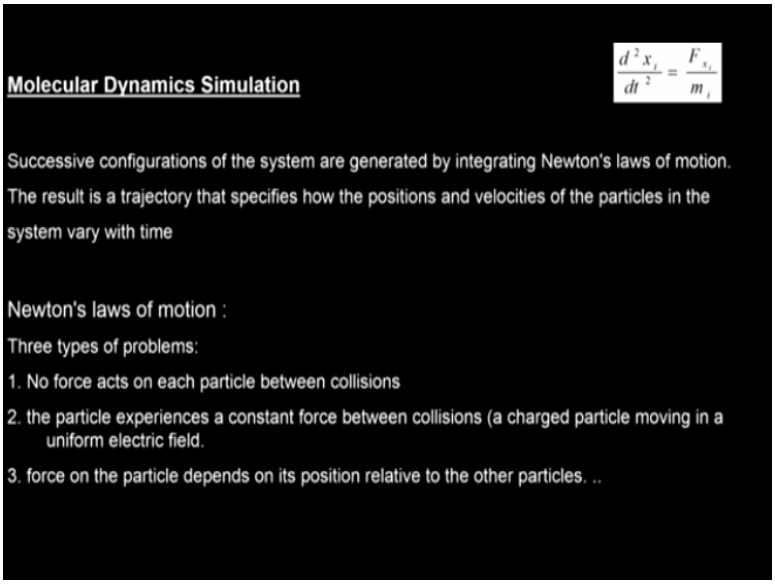


Once we do the conformation, we need to know whether molecule has attained is what is predicted in realities. As I had mentioned before X-ray is a very good technique, it gives you the coordinates of the molecule so we can cross check whether conformation matches Electron Diffraction, Raman's spectroscopy, Ultraviolet spectroscopy, IR, NMR, Microwave, Photoelectron, Supersonic Molecular Jet spectroscopy, Optical Rotatory Dispersion okay Quantum Mechanics so many.

And Quantum mechanics is theoretical approach. So, so many experimental approaches that can give you an idea about the location of the functional groups, location of the carbons, location of various atoms and then you cross check your predictions with the reality okay. So, so many different methods experimental methods and of course this is a quantum mechanics, this is a theoretical method to cross check whether you have reached the minimum energy conformation okay.

As I had mentioned that many cases, minimum energy conformation is not biologically active conformation and that might not be the real conformation when molecule goes and binds to the active side. So, you always try to look at conformations which are closer to minimum energy as well okay.

(Refer Slide Time: 20:24)



Another approach which can do is using molecular dynamic simulation. So, what you do in molecular dynamics, we need to basically you are solving this equation. This is Newton's second law d square x xi dt square = Force divided by m okay. Force = ma, you must have studied long time back. This is = acceleration. So, if you integrate this equation, you can get new coordinates starting from one set of coordinates at certain time point and given a force and given a mass.

So, you get successive conformation of the system as a function of time by solving this particular equation okay. So, it gives a trajectory that specifies how the position and velocity of the particle in the system vary with time okay. As you solve, you know the coordinates xi at a time at i. So, when you do ti + delta t you will find a new xi + delta t and so on actually. So basically you are solving that Newton's law of motion. 3 types of problems.

No force acts on each particle between collisions okay. The particle experiences a constant force between collisions that is charged particle moving in an uniform electric field. So, there is an uniform electric field so they uniformly feel a force, whereas the first one that is no force acting until they collide. Force on the particle depends on its position relate to the other particles okay. So there could be attraction, interpretations and so on.

So, these are different types of things that can happen when you are doing molecular dynamics. So, what you do, as I said integration is broken down into many small stages. Each separated in time by a fixed time delta t. basically it is a numerical integration. The total force on each particle in the configuration at time t is calculated as a vector. Some of its

interactions with other particles okay then from the force we can determine the accelerations then we get the positions and velocities at t + delta t.

That is what molecular dynamics does actually. The force is assumed to be constant during this time stamp. You cannot assume that, during this time stamp delta t, the force changes that complicates the whole issue actually. Then you determine the force of the particle in the new position okay. So, you again calculate t + 2 delta t positions at 2 delta t and so on. Velocities also we calculate.

At t + 2 delta t, in that way you keep on going forward so molecular dynamics is basically a forward starting from at t = t0, you run the simulation for certain pico seconds and then calculate the new set of coordinates of the molecule. So, all the algorithms assume that the position and dynamic properties that is the velocities, acceleration etc can be approximated at Taylor series okay.

Because, as I had mentioned solving differential equation is basically a Taylor series type of expansion okay. So, once you bring all the molecules together we sort of equilibrate these molecules so that they are not overlapping. Atoms do not overlap, they are sufficiently at for their distance and so on actually okay. So, that is what we try to do when we do the molecular dynamics.

So, these are the different types of techniques that are available for predicting or for telling whether you have reached the minimum energy conformation of the molecule okay. So, systematic search method, then of course problem becomes very difficult as the number of rotatable bonds become more, random search, you randomly go here, there here there and then try to see.

I mean here there means you get different conformations and see how the energy looks like Monte Carlo method based on statistical mechanics, genetic algorithm method is based on selecting those conformations which are favorable, rejecting and then coming up new set of conformation, almost like evolution coming up with new set of conformation based on the population of old set of conformation.

Then, of course we have the molecular dynamics approach where we start with one conformation at time = t0 and then you solve the Newton's second law force = ma is the acceleration that is d square x/dt square, where x is the coordinates of each of the atom and you go move from one-time step t0 to t0 + delta t, t 0 + 2 delta t and so on. And then calculate the new positions by solving these equations so the atoms move as a function of time.

So, there could be a constant force acting on them. For example, in a constant electric field, if the atoms are charged or there might not be any force acting when they are not colliding. When they are colliding of course because, of the collision there are some forces. So, you keep doing the solving this numerical differential equation from one-time point to another time point and so on.

So, when you start going up to large time steps, the molecule would have taken different conformations, which are very novel and then you can look at those conformations and see whether you from there minimize the energy and then see whether it has reached the minimum energy conformations. So, these are different approaches by which you can try to arrive at the minimum energy conformation.

And as I said, the minimum energy conformation need not be the conformation which the biological system may be reacting and sometimes, like I showed many examples the conformation between the minimum most energy system and those closer to minimum energy might not be much. So, you will find molecules with different conformation based on the Boltzmann equation e power -G/RT okay. So, the distributions will be in that order.

So, it is always good to look at the conformations of molecule which are slightly above the minimum energy conformation also okay. So, with that we complete the topic of molecular mechanic’s force fields. So, because this is the most important, we had spent lot of time on this and as I showed you a free software, which can get you to draw structures of organic molecules or we can import structures from for example hexane or even from zinc and then we can minimize energy.

It has got only one or 2 couple of force fields. But, if you look at other software's, you may find more force fields. So, we can look at conformations and then after that we can do lot of

property calculations okay. So, with that we will stop and in the next class we will talk about quantum mechanics okay. Thank you very much for your time.