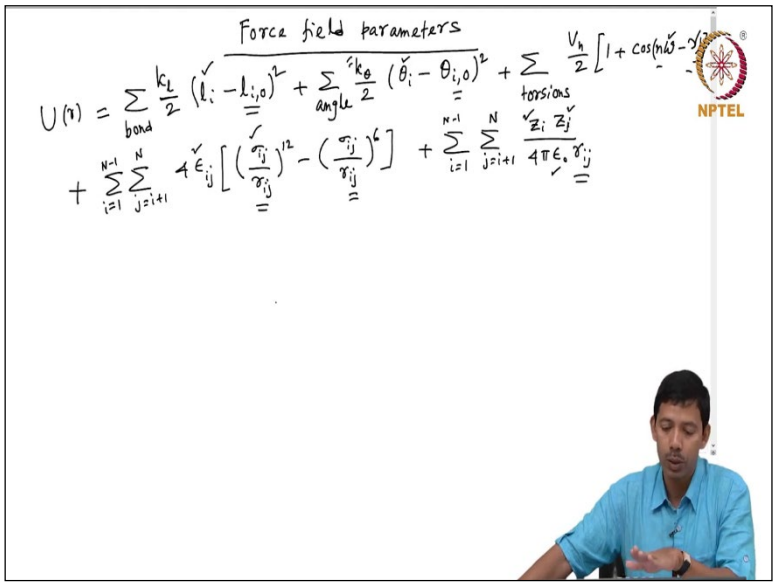
Thermodynamics for Biological Systems:
Classical and Statistical Aspects
Prof. Sanjib Senapati
Department of Biotechnology
Indian institute of Technology - Madras

Lecture – 79
Understanding Force Fields

(Refer Slide Time: 00:26)



So today I will be talking about more about the force field parameters. I have introduced you the force field before but we need to know a little more about it. So, to talk about force field parameters I need to write down the total potential once again for a molecular system.

$$U(r) = \sum_{bond} \frac{k_l}{2}(l_i - l_{i,0})^2 + \sum_{angle} \frac{k_\theta}{2}(\theta_i - \theta_{i,0})^2 + \sum_{torsions} \frac{V_n}{2}[1 + \cos(n\omega - \gamma)] + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] - \sum_{i=1}^{N-1}\sum_{i=1}^{N-1}\frac{Z_i Z_j}{4\pi\varepsilon_o r_{ij}}$$

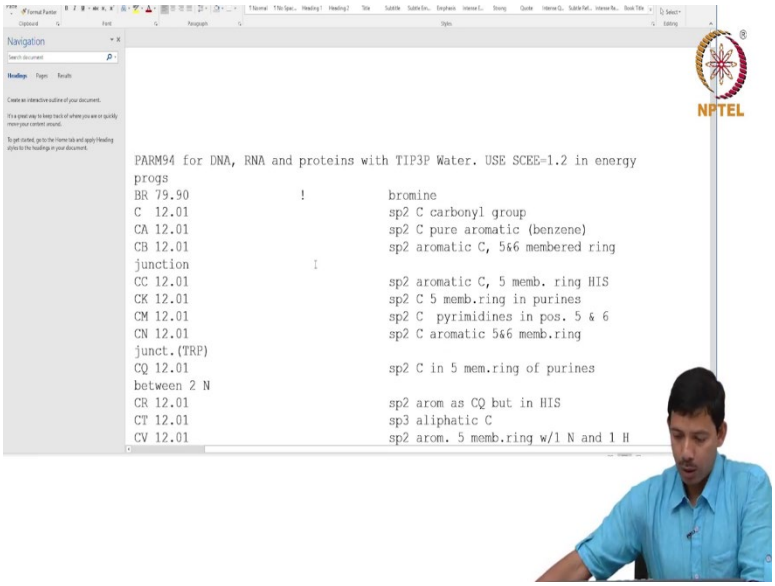
Here, U is the total potential of a molecular system, k_l is the force constant, l_i is the instantaneous value of bond length, $l_{i,0}$ is the equilibrium bond length of ith bond, k_θ is the force constant, θ_i is the instantaneous value of angle, $\theta_{i,0}$ is the equilibrium value of angle, V_n is the barrier height, n is multiplicity, ω is the instantaneous value of the dihedral and γ is the phase factor. Till here the terms represent the intramolecular interactions. This is followed by two intermolecular interactions- Vander Waal which is Lennard-Jones and electrostatic terms.

$(\sigma_{ij}/r_{ij})^{12}$ represents repulsion between ith and jth particles, $(\sigma_{ij}/r_{ij})^{6}$ represents attraction between ith and jth particles. Zi is the charge on ith particle, Zj is charged on jth particle, ε_0 is the dielectric constant of the medium, r_{ij} is the distance between ith and jth particles.

So, this was the total of potential of a molecular system. Now in this expression l_i is a variable which is the instantaneous value of the bond, θ_i is a variable which is the instantaneous value of the angle. ω is the instantaneous value of the dihedral and r_{ij} is the variable which is the inter particle distance between particle i and j. Everything else in these expressions are constant.

So the equilibrium parameters they remain fixed therefore we can put all this equilibrium parameters together and make a library. If you have this library for all sorts of possible molecules then I do not have to find those equilibrium parameters here and there I just go to that library and take off the required parameters from that particular library. And based on the available parameters scientists have made such libraries one such library.

(Refer Slide Time: 06:21)



That library is called AMBER PARM 94 parameters in AMBER package. So if you look at this PARM 94 parameters which is part of MD simulation software called Amber, it lists out all these equilibrium parameters often from various molecules. So, at the top of this parameter which are the force field parameters you see it shows the name of the atom and its molecular weight. So, here is carbon and the carbon is 12, C alpha C beta 12, so carbon can be of different types this C is sp2 carbonyl carbon C alpha CA is sp2 carbon in pure aromatic like benzene.

So all possible carbons are listed up here. And then you have calcium you have fluorine you have hydrogen of different types and the molecular weight is given in the next column and also for various other atoms iodine, chloride, sodium, magnesium, different types of nitrogen and so on so forth. There are many other atom types possible and those atom types and their molecular weight is listed at the top of this library in the force field parameter file.

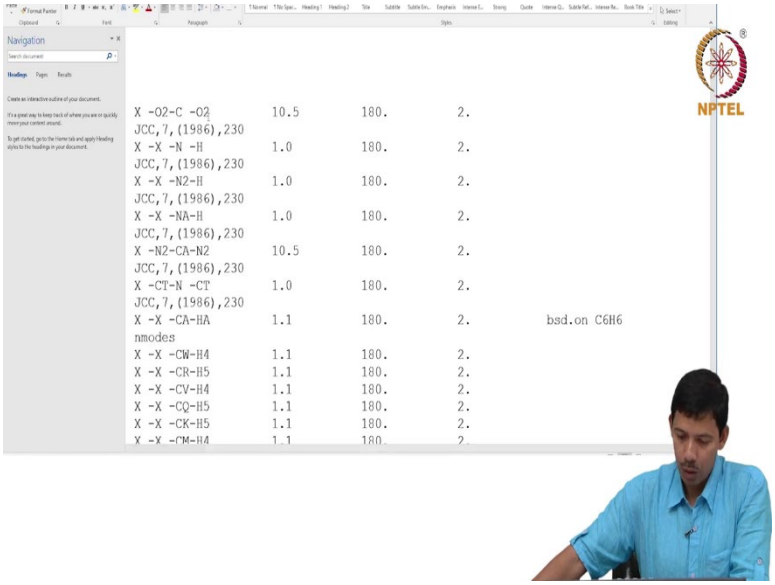
Next comes the bond so if you recall the bond term, in the bond term you had the force constant and then you have the equilibrium bond length and here exactly those are listed so the first one is basically water oxygen, water hydrogen. So for that particular bond what is the force constant and what is the equilibrium bond length. So all these bond lengths are given in Å and the energy values are given in kcal/mole.

So as you see here for water oxygen hydrogen, the bond distance is 0.9572 for water model called TIP3p and the 553 is the force constant. The distance between hydrogen 1 and hydrogen 2 is 1.5136. Different type of possible bonds are listed here with their force constant and their equilibrium bond distance in this library file.

If you go further down then comes the angle. So, here in the angle list the first one is hydrogen of water, oxygen of water, with hydrogen of water so this HOH is having a force constant of 100 with 104.52 as an equilibrium angle of TIP3p water. Likewise all kinds of possible angles are listed here.

You have all kinds of angles listed here I am just looking for something like here you have a CT which is basically the sp3 carbon carbon so that is 40 force constant and 109.50 is a equilibrium angle. So, all possible angles are listed in here. After bond and angle we now can look for the dihedral.

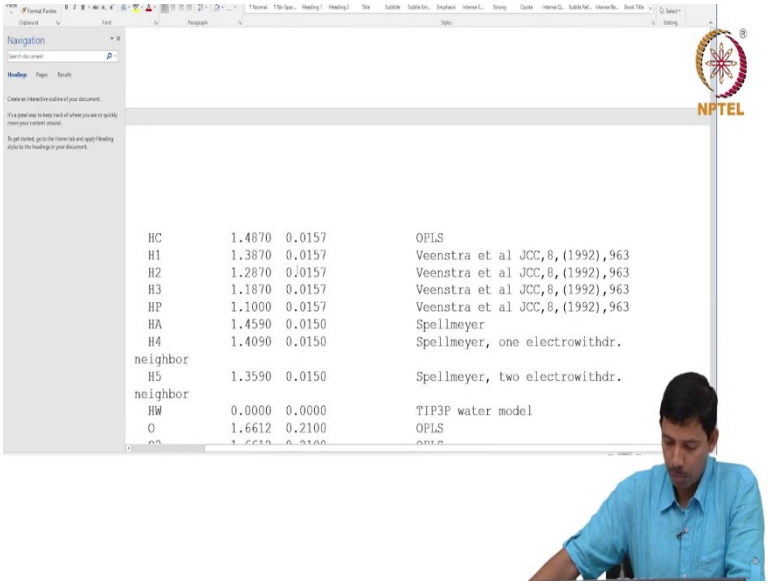
(Refer Slide Time: 11:16)



So in the dihedral, you can see that for the dihedral 1 2 3 4 if you recall the dihedral is having 4 atoms four connected atoms so here for this dihedral here X meaning it can be any atom so here any atom, carbon, hydrogen, oxygen, nitrogen and so on and so forth. So, for any atoms 1.1 is a barrier height 180 is the phase factor and 2 is the multiplicity.

So here for many possible dihedrals the Vn phase factor and the multiplicity are given. So, all possible dihedrals are listed up here.

(Refer Slide Time: 12:45)



If we come further down then we have the ϵ and σ parameters. So, this is the σ parameter for of the Lennard-Jones and this is the ϵ parameter of Lennard Jones. So, in Amber this ϵ is multiplied by a factor you should look at them when you are going to use these particular force field

parameters. Nevertheless this is a library available in Amber software which you can use if you run your biomolecular system or liquid system in Amber.

And you can just make use of this library, from where the algorithm will pick up your bond angle or dihedral of interest and make use of them for the calculation. Now so if you recall that we talked about a model system when we try to simulate we need to make a model system and when you make a model system we basically start making a model and then iterate it until we reach to certain experimental properties.

For example I take the simple example of liquid water. So the liquid waters density 0.998 gm/cc. For water the diffusivity is known density is known potential energy is also known. So, if I use this bond angle and the Lennard Jones ϵ, σ parameters for water and then carry out a simulation I could easily get the density and diffusivity and potential energy. so, now I match those values with the experimentally known parameters and see how good or how bad my match is. If my match is not good so I will go back and tune my bond, angle, epsilon, Sigma parameters and rerun the simulation until I will get good match of my water properties with the experimentally known properties.

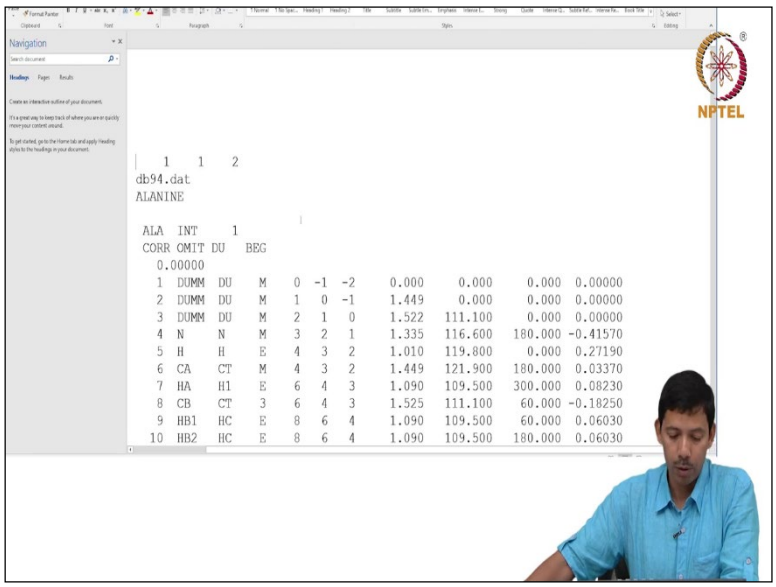
Once I get a good match then I know that this set of parameters are really good. So, then I put up all these parameters I can modify these parameters with the permission of the owner of the library in amber group and then I can remake this library with those parameters which reproduces the experimental data and that is how all these parameters are developed and being deposited and that is why these parameters are also called transferable parameters.

So, the meaning of transferable parameter is basically even though these parameters are obtained from bio molecular point of view from given sequences of proteins, since my 20 amino acids are same irrespective of protein A to protein B so these parameters what we have generated looking at the experimental properties of known proteins can be applied for an unknown sequence or a protein for which you know you have a sequence and the structure is not known.

So, I can easily make use of these parameters and use those parameters for the other sequence what I have and predict three dimensional structure of that protein and that is the meaning of the transferability of these parameters. So, one thing you might have noticed that in this library everything else is there except my except my Zi and Zj the partial charges on ith particle and jth

particle ε_0 is a dielectric constant of the medium which is 78 if you if your medium is water. So that is a fixed quantity now everything else we have got in the library but not Zi and Zj .

(Refer Slide Time: 17:24)



So for that now there is another library which. This is also part of the Amber software. And if you look here, in this library file we do have all 20 amino acids and all their parameters are written out, like here for alanine the very last column is basically depicting the charges of each of the alanine atoms. So, alanine is having peptide nitrogen, the connected hydrogens, C alpha H alpha C beta with beta you have H B1 H B2 hydrogen etc.

And this is carbonyl carbon-carbonyl oxygen and for each of them the partial charges are listed in the last column these 3 columns are basically giving the bond angle and dihedral. So, this bond angle dihedral and the charges on the atoms of alanine and for that matter all other amino acids we have a file where we have all the bond angle dihedral and also the charge is written.

So again these parameters are obtained as I said by looking at various proteins of known structures and based on their structures scientists have generated these parameters and then they put up in this library. And now if you use a different sequence of glycine alanine lysine and so on so forth, all these parameters would hold good because the same 20 amino acids are present in protein A versus protein B.

Now here what you see is the way for the 20 amino acids, the bond angle dihedral are written in a format called a Z matrix. For a molecule you can basically write the structure of a molecule in a Z matrix form and that is how it is written for these amino acids. So, what is Z matrix? So, Z

matrix basically puts up the structure of a molecule in terms of its internal coordinates. So, what are the internal coordinates? Internal coordinates are basically the bond angle and dihedral.

So bond angles and dihedral they make the structure of a molecule, the internal structure of a molecule and therefore Z matrix is often used for a molecule whose structures are well known. So, since we know the structures of amino acids very well so we can write the structures in terms of Z matrix. The advantage of Z matrix over writing XYZ coordinates for each atom is that based on XYZ coordinates I do not have to calculate bond and the dihedral separately.

I have bond angle already calculated for these 20 known amino acids and therefore my computational cost will be lowered that means the calculation will be faster if I have the Z matrix form of a molecule already written rather than their XYZ coordinates. So, now I will explain you what is Z matrix and how we basically make our Z matrix? How this Z matrix of Threonine is made?

So now I will talk about the Z matrix but I hope that the force field parameters and how they are developed why they are useful must be clear to you because they are very important for generating the new microstates starting from microstate one. Basically you have a microstate from initial coordinates from PDB or homology model and then you basically evolve. How we evolve from one parameter to the other parameters one structure to other structures that I will explain in further lectures.

But to evolve from one structure to another structure you need force field parameters and these force field parameters are very important and are available in literature.