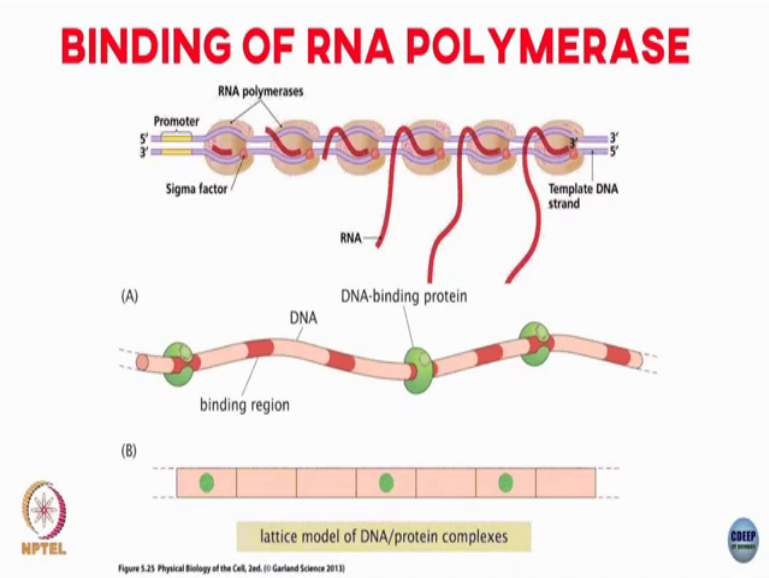
Physics of Biological Systems
Prof. Mithun Mitra
Department of Physics
Indian Institute of Technology, Bombay

Lecture – 24
Binding problems

(Refer Slide Time: 00:17)

So, it is this entropy that will be concerning ourselves with today.

(Refer Slide Time: 00:21)

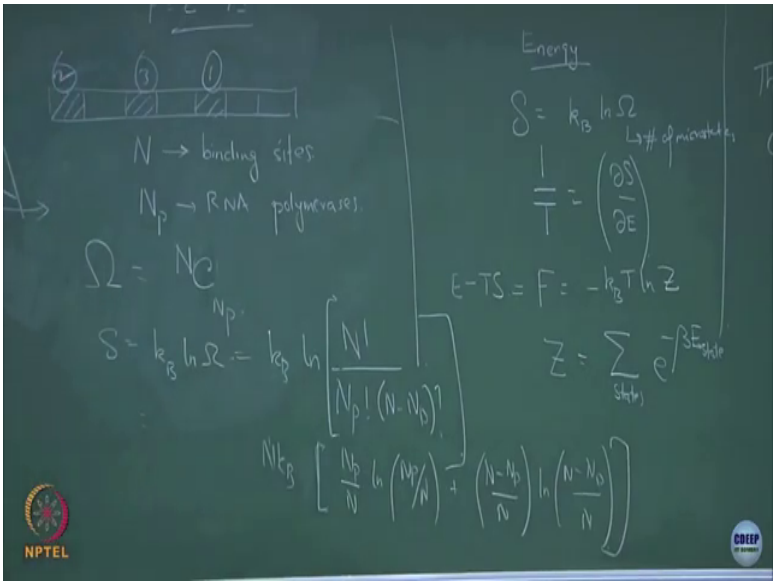


There is a very simple calculation to get started on. So, here is RNA polymerase this example that we were talking about today. This RNA polymerase binds on the DNA it translates along this template DNA strand and it produces the same RNA, ok. So, if you again, if I am just concerned about the binding of this RNA polymerase to the DNA, I can aprox I so, I create a cartoon in my mind of this process that this DNA has certain regions which are binding regions for the RNA polymerase, right.

So, given by this dark red regions, so, if I think of this DNA as a 1 dimensional lattice, there are certain sites where the RNA polymerase can come and bind. This is true not only for RNA polymerase, but for any DNA binding protein. So, it is a genetic model for a generic lattice model for a DNA protein complex. So, in RNA polymerase or it is some protein can come and bind on certain specific sites. And maybe you want to calculate what is the probability of

occupy. So, what fraction of these binding sites on the DNA will be occupied by an RNA polymerase.

(Refer Slide Time: 01:52)



We will do that, but let us first do a simpler problem that what is the let us say I have some. So, remember the free energy is E minus TS. So, let me for the time being first consider just the center of the part I will come back to the energy part later on. Let us say I have this 1 dimensional DNA this is of course an approximation in that I am forgetting about the sequence specificity and so on. I am just saying that certain sites can bind my RNA polymerase.

So, let us say I have N sites that can that are binding sites and I have N P of these RNA polymerases, right. What would be the; what would be the number of microstates? So, in how many ways can I arrange these N P molecules among these N available binding sites? Yes N. Why a permutation?

Student: (Refer Time: 03:28).

So, for example, almost correct. So, let us say I have an RNA polymerase, I put the first RNA polymerase somewhere here, I put the second one over there 1 2 I put the 3rd one here, right. This is indistinguishable from the fact that if I had put the first one here and maybe the second one there right, because this is RNA polymerase it is difficult to distinguish which each individual RNA polymer.

All you can see is whether in the site is occupied by an RNA polymerase or not which means that this p is not the right thing, but it becomes C, right. And then the entropy is of course, k B log of omega, right. Is this clear, right? If you have N sites in how many ways can you choose to fill N P of them? What is this entropy, anyone? This is the simple to state problem, right. If you remember your classical statistical mechanics, you have some fill sites you have some empty site. What is the corresponding entropy, but even if you do not remember the expression how does this entropy behave.

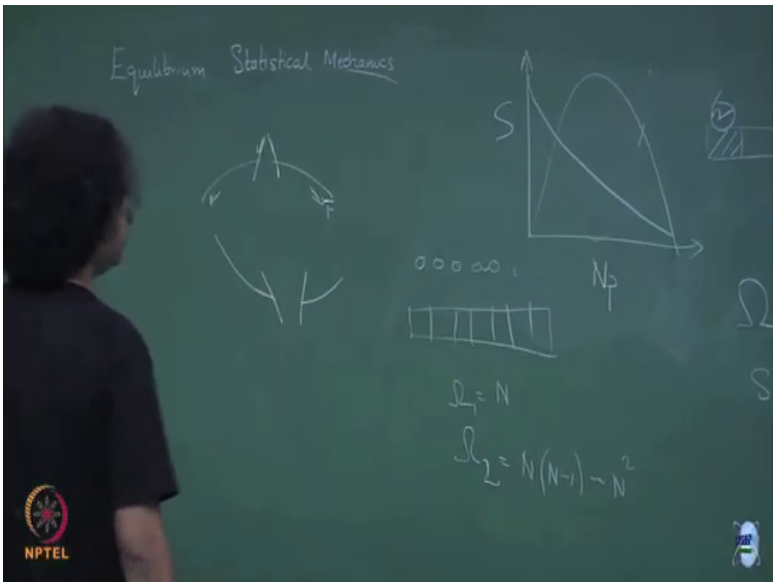
As let us say a function of P. If I were to plot this entropy as a function of N P. So, I can calculate this of course, I can do k B log of N factorial by N P factorial N minus N P factorial, right. You use the Sterling's approximation, you assume that N is large and P is large and so on. And you can reduce this hopefully to a recognizable form, but this is fairly elementary a two state system in stat mech.

Anyone remembers how the entropy looks like how those entropy behaves? Maximum at the center where as some (Refer Time: 05:48) is saying like this. Then it will fall like something like this, ok. Sahithya, tell me why a maximum at the center.

Student: (Refer Time: 06:15).

When N P is N by 2, ok. Shogata, why will it decrease?

(Refer Slide Time: 06:32)

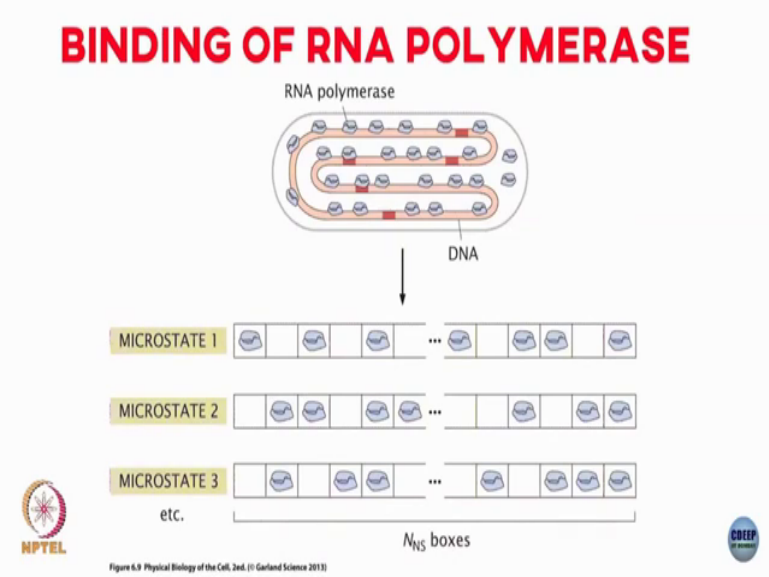


Let us say I had these binding sites and I wanted to I had only one RNA polymerase, right. So, this one polymerase I can place it any one of these N ways. So, my omega each of them is a configuration. So, my omega is N, right. Say, if I had 2. So, this is omega 1 number of states. So, if I had omega T 2, the first one I could have placed N ways, the next one I could have placed in N minus 1 ways, it was something like N squared, right.

So, if I take the log of this which one is going to be larger, right. So, similarly if you had many, many, many, many, RNA polymerases. Let us say you had equal number of RNA polymerases where you had sites, you again come back to very few options, because if you had equal exactly equally you had only one way which is everything is filled, ok. So, this is the right, this is the right behavior of entropy. It is very low when you have no molecules, it is

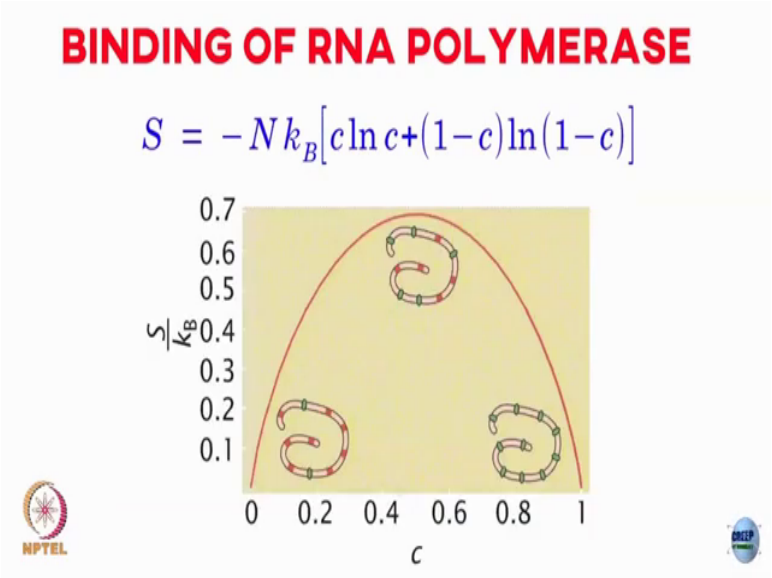
very low when you have very high molecules, it is exactly its maximum exactly at N P equal to N by 2.

(Refer Slide Time: 08:15)



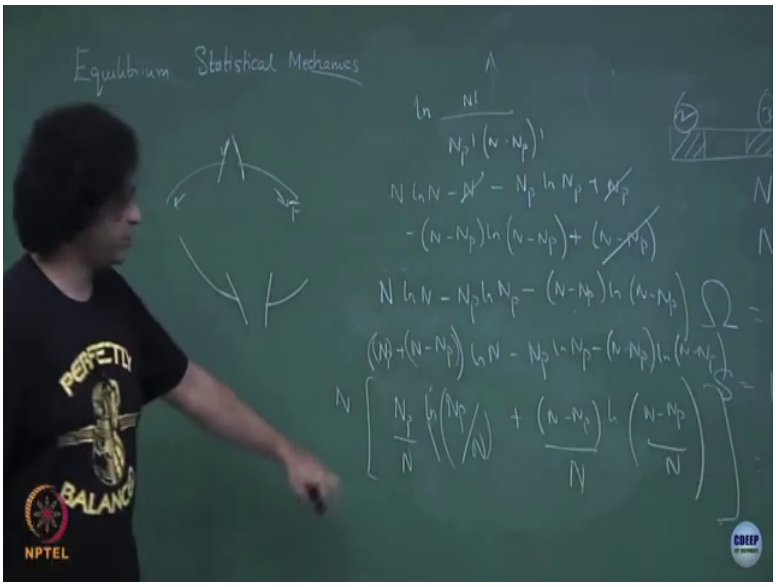
So, this is the this is what is this is the microstate sort of counting which we have done.

(Refer Slide Time: 08:22)



Now, anyone does that make you recall the formula for the entropy no where is my. So, you should again go back and recall. So, if you do this Stirling's approximation on this what you will get is something like N k B, right log of N P by N, N P by N log of N P by N plus 1 N minus N P by N log of N minus N P, right. Does this make sense?

(Refer Slide Time: 09:21)



Actually we might as well this (Refer Time: 0:09:19). This is log of N factorial by N P factorial N minus N P factorial which is N log N minus N that is this minus N P log N P plus N P. And then minus N minus N P log N minus N P plus N minus N P this is Stirling's approximation.

So, this n will cancel with this N p and N minus N P. So, what you are left with is N log N minus N P log N P minus N minus N P log of N minus N P, right and then you just massage it. So, you can break this N up into N P plus N minus N P, right log N and then the remaining two terms you keep the same. And then if you divide this so, see you take N P common then it becomes log N minus log N P.

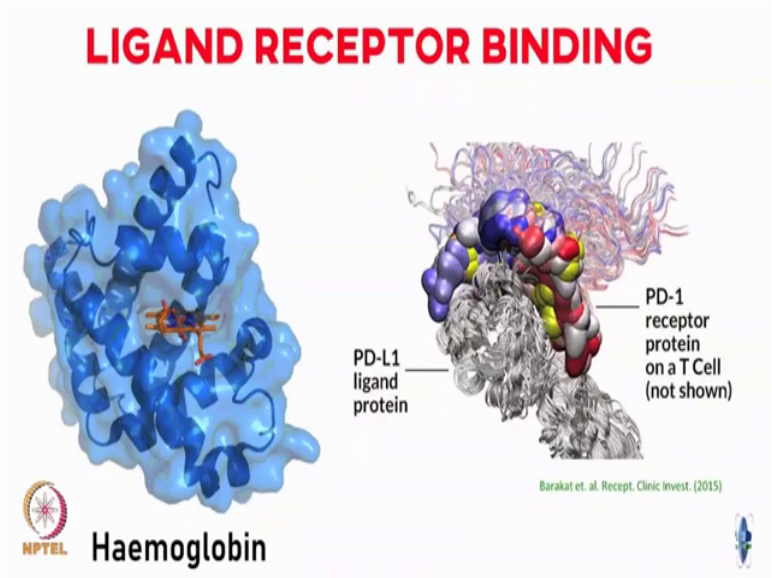
So, this is N P log N P by N, and then, so this N minus N P common so, plus N minus N P again log of N minus N P by N, right. And then if you take an overall N outside then you

divide by N log and this is a standard two state entropy, right. So, if in some sense you call this the linear density of polymerase molecules or whatever molecules that you are interested in if I call this quantity C if I call this like the C, then this is C log C plus 1 minus C log 1 minus C, right which is what we have.

So, N k B c log c plus 1 minus c log 1 c. And that looks if you plot that entropy that will look exactly like this the maximum number of microstates that you will have available is when N P is exactly half, half of N so, c is equal to half that is the standard to state entropy.

So, you can use these sort of basics (Refer Time:12:02) concepts to get a handle on what are the entropy or what are the number of microstates that a binding process like this would have in a biological context. We have of course, done many simplifications, we have not taken any specificities, we have forgotten about the actual 3 D structure of the DNA and so on, but even so, even with these we will see how good these approximation these type of simple calculations too, ok.

(Refer Slide Time: 12:31)

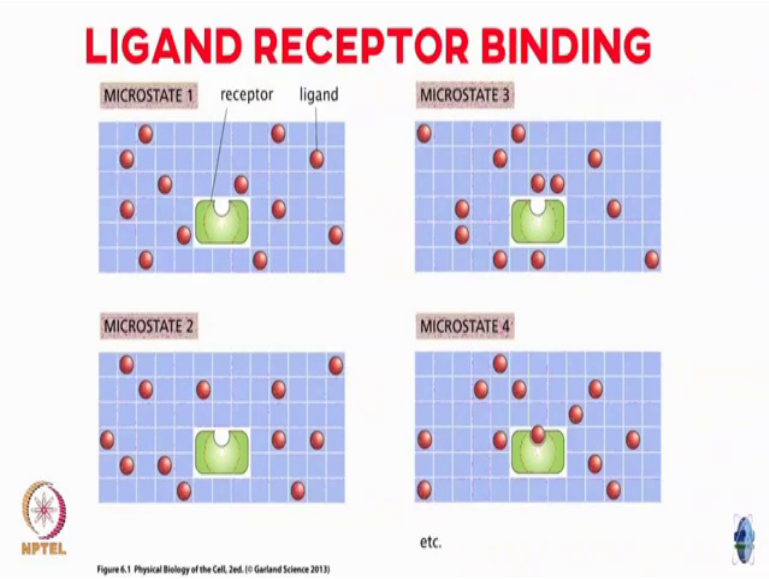


I will come back to the, but let us look at for the moment related problem of this ligand receptor binding, ok. Again this is generically of the same class of problems that we are thinking about it is a binding problem. So, I have two protein molecules, ligand and a receptor.

For example, this is the haemoglobin which binds this oxygen and the haem group at the center binds the oxygen so the oxygen would be the ligand, the haemoglobin protein would be the receptor. Similarly, this is a different protein molecule, this is the PD 1 ligand protein, this thing over here this curve protein over there is the receptor, ok. So, you have a large protein generally which we call as the receptor and a smaller protein which comes and binds on it which we call the ligand.

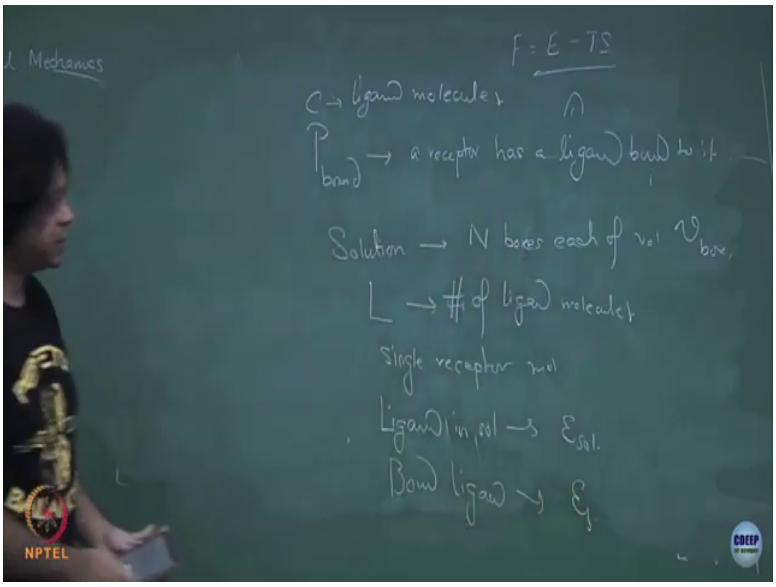
So, I have some large protein let us say like this which is my ligand and a smaller protein comes and sits on it which is my receptor, ok. And then depending on whether this protein is bound or not you can perform different functions inside the cell.

(Refer Slide Time: 13:38)



So, again let us say what I want to ask is that: given a certain concentration of these ligands, what is the probability that a receptor is bound to alike?

(Refer Slide Time: 13:58)



So, let us say, let me define the problem. So, let us say I have a certain concentration of ligands, certain concentration of ligand molecules. And I want to ask that what is the probability let us call it P bound that receptor will have a ligand bound to receptor has a ligand bound it. So, what we will do is again we will make a very simple approximation, we will consider this sort of a lattice model, ok. So, here is my receptor molecule that is sitting over here and I have these ligands floating about in solution, ok.

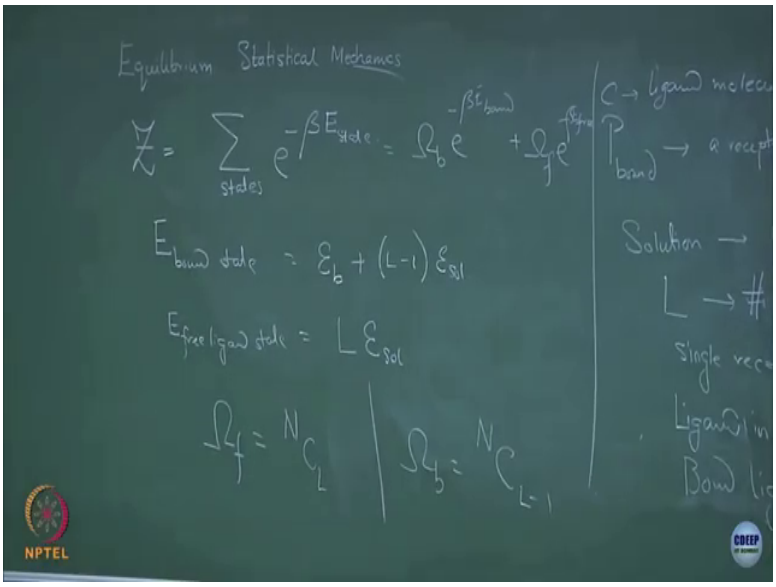
So, we will consider a lattice model of the solution. Basically, I will say that each microscopic volume is something let us call some V box and these ligands will occupy one of these boxes, ok each ligand will occupy one box and then we will try to calculate the entropy and from there, we will try to calculate what is the probability that this receptor will be bound to the ligand, ok so let me first define again.

This, so my solution for the solution, I will approximate it as a lattice model. So, N boxes each of volume let us call it some V box, that is why model for the solution. Let us say I have L number of ligand molecules number of ligand molecules, ok and I have a single receptor molecule the single receptor molecule so exactly like that cartoon, ok.

Now, first when a ligand binds to the receptor it sort of lowers its energy, it is energetically favorable for this ligand to come and bind to the receptor right. So, let us say a ligand in solution has some energy ligand in solution; ligand in solution has some energy, let me call it epsilon solution and like a bound ligand has some bound ligand has a different energy let us call it epsilon bound which is lower than this epsilon solution.

So, it lowers its energy by binding to the ligand. So, that is what it would like to do energetically, but again as we discussed entropically, it is going to be favorable for it not to stay bound, it would like to stay in the solution. So, energetically it is favorable to be bound entropically it is favorable to be unbound. And then what it will actually be or what is the probability of finding a bound the receptor with a bound ligand will depend on this competition between the energy and the entropy ok this clear the at least the setup is clear right.

(Refer Slide Time: 17:14)



So, what we will first do is that we will calculate the partition function. Remember, the partition function is the sum over all states E to the power Boltzmann factor E to the power of minus beta times the energy of that state, ok. Now, let me say broadly there are two states: one state in which the ligand is bound to a receptor and another state in which the ligand is unbound. So, let me say the energy of the bound, you will have an energy of the bound state and an energy of a free ligand state, ok.

So, if you have a ligand bound to the receptor, what will they be the energy of the system as a whole? Yes or let us say let us talk about this first if no ligand is bound to the receptor what would be the energy?

Student: L into (Refer Time: 18:26).

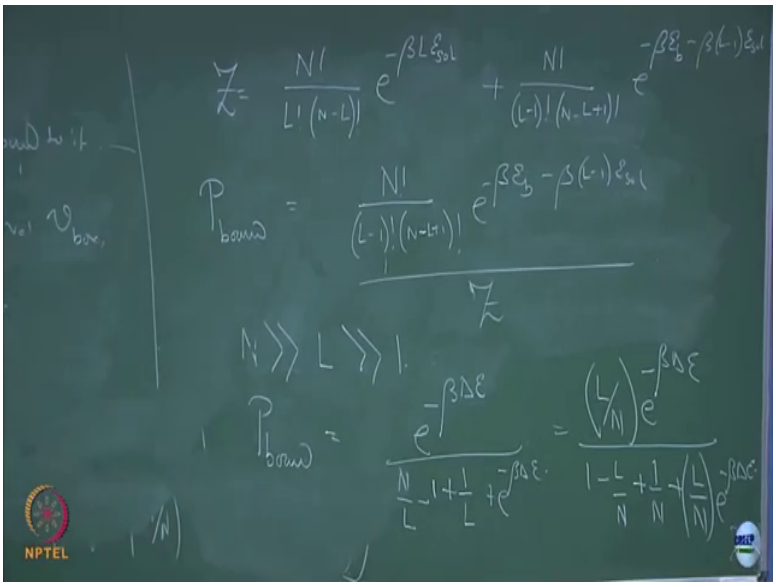
L into epsilon solution, right. All my ligands are free in solution I have L of these, so L into epsilon solution. Now, if I have this bound state, but one of these ligands is bound to the receptor what will be the energy of that state?

Epsilon bound plus L minus 1 Epsilon sol, right. One of these ligands is bound the remaining L minus 1 are in solution, ok. What is at least I know these two states ah, but there are many ways to get these states right. So, if I consider in terms of these bound state and the free state then I could say that what are the multiplicities of these states, right. What is omega b e to the power of minus beta e bound plus omega free e to the power of minus beta e free, right. Where the c bound and e free we have calculated. So, it is left to calculate is this omega b and this omega f, ok.

What is this omega b going to be or what is this omega f going to be? So, in this omega f I have these L ligands, right which I have to or arrange among these N available boxes of the solution. So, that is N C L, right N C L and that is why and therefore, that will give me the entropy that my point.

What would be omega b? N C L minus 1, right 1 is bound so that does not have any freedom anymore, you have the remaining L minus 1 ligands which are still in solution therefore, N C L minus 1, good.

(Refer Slide Time: 20:46)



So, once I have that I have this partition function in principle, so my partition function Z is N factorial; N factorial by L factorial N minus L factorial that is the free so then e to the power of minus beta L epsilon sol plus N factorial by L minus 1 factorial N minus L plus 1 factorial e to the power of minus beta E b minus beta into L minus 1 epsilon sol ok, right. This is my full partition function, yes.

Student: (Refer Time: 21:36).

Yes, well not really right, because that would not change between these two states between the bound state and the free state, you would have some factor of course, which would encapsulate the receptor mobility but that would be the same here versus here right. So,

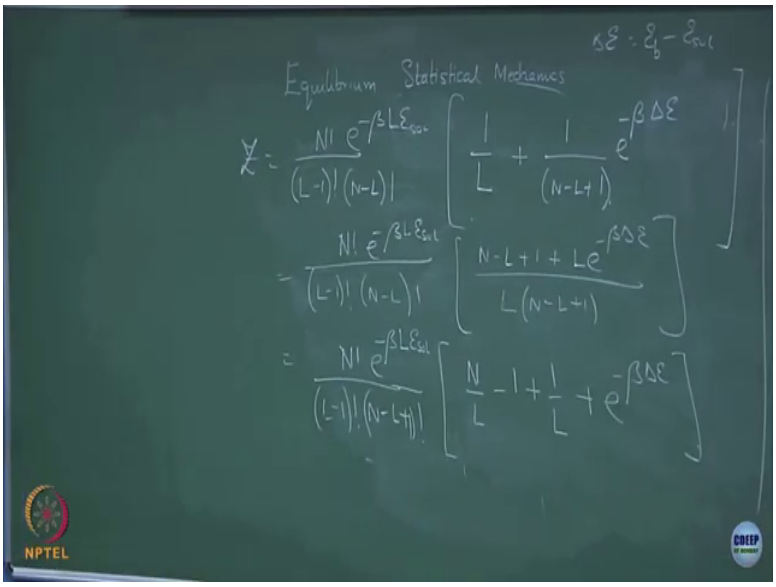
effectively it would not make a difference as far as this type this question is concerned, ok so to using, so this is my partition function.

Now, if I am interested remember, when I am interested in what is the probability that the receptor has a ligand bound to it. So, if I ask what is P bound, now that I know the partition function, can you tell me what is the probability the receptor has the ligand bound to it? Yes Sowmya? Which one by which one? First one, the second one divided by the partition function these are all the states that have a receptor bound to a ligand, right.

Therefore, the probability that you have a receptor bound with the ligand with this N factorial by L minus 1 factorial N minus L plus 1 factorial e factorial e to the power of minus beta epsilon b minus beta L minus epsilon sol divided by that whole thing divided by Z, right. Now, this looks bad right, but we can make a few we can try to simplify this by using Stirling's and sum the.

So, let us say we are working in the lim we are working in a sort of thermodynamic limit where N is much much larger than L, but L is also a much much larger than 1, ok you have lots and lots of ligands, but equivalently, you have lots of much more of free space as well, right. So, N much much greater than L much much greater than 1. So, for both of these N and L terms I can use the Stirling's approximation so that is what I will do. So, let me try to use it on, how much time to it I have time. So, let me try to use it on the partition function, let me see.

(Refer Slide Time: 24:00)



So Z is for that expression that I have written there. So, let me say I take N factorial is common, L minus 1 factorial is common and then N minus L factorial is common, right N minus L factorial is common. And then e to the power of minus beta e to the power of minus beta L minus 1 epsilon sol, ok. It will let me just take beta L and (Refer Time: 24:45) just convert that beta L eps epsilon sol, ok.

So, then in the first one in the first term of the partition function all I have left is a 1 over L, right L into L minus 1 will make give me that L factorial. In the second term what I have is 1 by N minus L plus 1, right which into this will give me a N minus L plus 1 factorial. And on top I have e to the power of minus beta epsilon b minus epsilon sol, right. So, that this and this together gives me the L minus 1 epsilon sol this is the minus beta epsilon b.

So, let me call this as some delta E this is the amount that you lower your free energy is that you lower your energy by ligand when ligand binds to the receptor. So, let me call this as my delta some delta. So, delta E is epsilon b minus epsilon sol. So, this is whatever is outside is outside. So, let me see (Refer Time: 25:35) is N minus L plus 1 plus L e to the power of minus beta delta E by L into N minus L plus 1, right. And what I want is so over here was N factorial e to the power of minus beta L epsilon sol by L minus 1 factorial N minus L factorial.

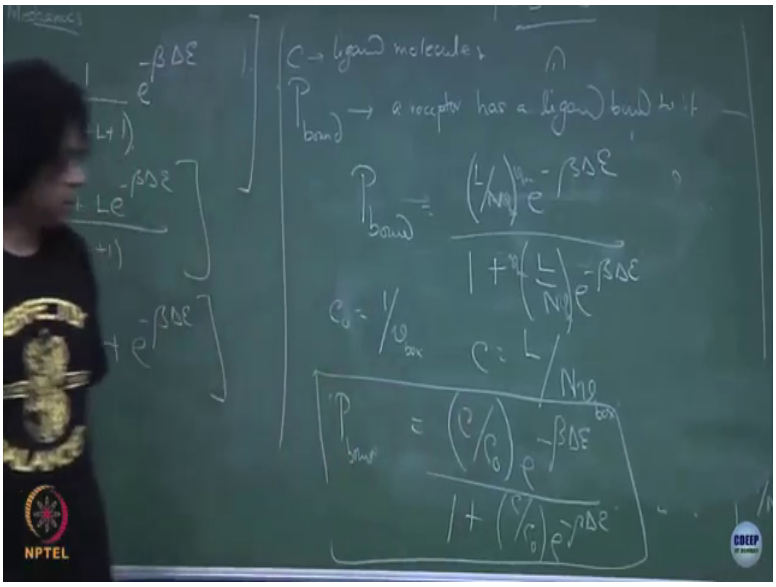
Let me pull this one outside so, N factorial by L minus 1 factorial N minus L plus 1 factorial. Then what I am left with is; N by L so N by L minus 1 plus 1 by L plus e to the power of minus beta delta E, I hope I have the signs, right. Why I wrote it in this form is that ultimately I want to put that Z over here. So, I just want to get the pre factor is the same which is N factorial L minus 1 factorial N minus L plus 1 factorial which is and of course, I am missing this e to the power of minus beta L epsilon sol ok.

So, now if I substitute, if I put this back over there in this P bound expression, I will get rid of this pre factor completely. So, P bound is going to be what? Here again I will get e to the power of minus beta delta E, right anything else, beta delta E. And on the bottom I will get whatever is over here so N by L; N by L minus 1 plus 1 by L plus e to the power of minus beta delta E ok.

This is just manipulated it, I have got rid of this pre factor completely, I have got rid of this epsilon sol and epsilon b separately and written everything as delta E. The difference in free energy; difference in energy is between ligand in solution and a bind left bound like that.

So, let me now multiply and divide by this term L over N so L by N e to the power of minus beta delta E 1 minus L by N into plus 1 by N plus L by N e to the power minus beta delta E, ok. Now, because I am working in this limit that N is much much greater than L is much much greater than 1, let me neglect these terms in comparison to 1, right L by N is a small quantity 1 by N is an even smaller quantity.

(Refer Slide Time: 30:36)



If I do this so, in this limit what I get is my P bound, there is a much simpler expression for this P bound the probability that I will have a ligand bound to the receptor which is L by N e to the power of minus beta delta E divided by 1 plus L by N e to the power of minus beta delta E ok. If you want to write in terms of concentration, the total volume of the box is of course, N times V box. Remember each a was so, if I multiply and divide by V box and V box, ok.

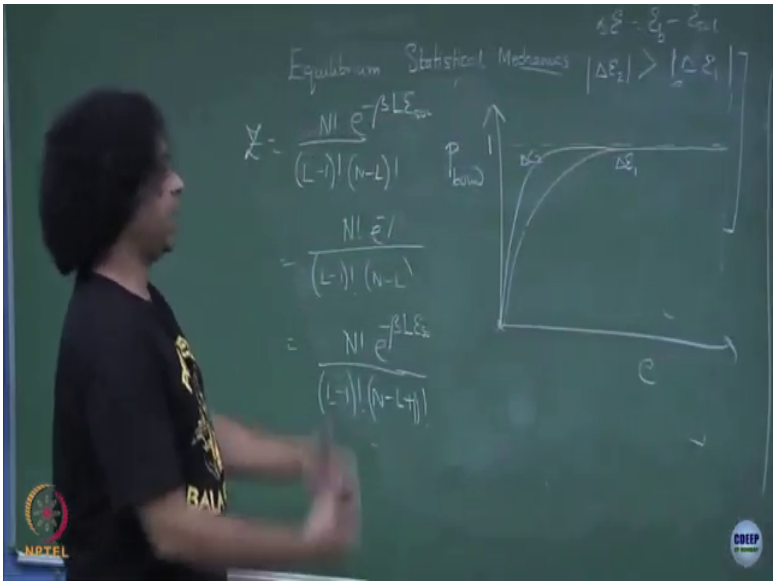
This is the concentration of ligands the number per in the whole volume and same here v box; v box. And let me say C, I call C naught as some concentration which is a reference concentration whatever that is. And then the this concentration of ligands C is nothing but the number of ligands in this whole solution N by v box, ok.

If I want to write in this concentration language, I might have left it here, but if I want since I asked as a function of concentration, if I want to write it in this concentration language this is

C by C naught e to the power of minus beta delta E by 1 plus C sorry, C by C naught e to the power of minus beta delta E.

So, given a certain concentration of ligands the probability that you will have a receptor which is bound to a ligand is given by this it depends on the concentration of course, but it also depends on how much free energy how much energy you lower by binding to the receptor molecule, right.

(Refer Slide Time: 32:35)



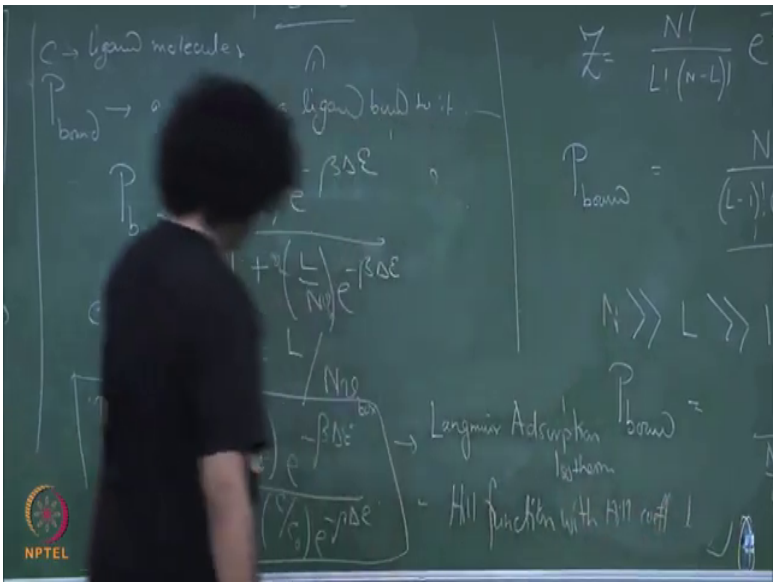
So, for example, if I were to plot it, if I wanted to plot this probability the probability that the ligand will be bound this P bound; as a function of the concentration of ligands, ok. So, if there was no ligand of course, then of course, the probability is 0, because there is nothing to bind so I know it will start off from here.

If the concentration was extremely high, then what would the limiting probability be? If the C went to infinity what would the probability that the receptor would be bound be 1. So, I know that at very high concentrations, it has to go to 1, right. And then you can plot this and see how this goes so it will go something like this, right.

This is for example, for some given value of delta E, let us say let me call it delta E 1, ok. If I now had a greater delta E, so for example, if this ligand receptor binding was even stronger, right it lowered its energy more by binding to the receptor. So, if I now wanted to plot it for a delta E 2 where this mod of delta E 2 is greater than mod of delta E 1. Which way would the curve shape? Would it come below or above? Above, right.

So, if I wanted to plot it for a delta E 2 then at the same concentration of ligands you would get a higher probability that the receptor is bound. This sort of a curve has a named, this is a very famous equation for binding it is called the language adsorption isotherm. It is called the language adsorption isotherm; adsorption isotherm.

(Refer Slide Time: 34:31)



This sort of a function x by 1 plus x it is also called the Hill function it is called a Hill function with a Hill coefficient of 1 with it I will explain with the coefficients. So generally, you can have hill functions or functions where you will have something like this raised to the power of N, ok.

In this case N is 1 which means this N is called the Hill coefficient which means is the Hill function with Hill coefficient 1. Later on, as we do as we continue to do problems like this, you will see that in cases where you have cooperative binding where the fact that you have bound 1 ligand means it is more likely that you can bind another for this multivalent receptors which can bind multiple things you will see that you get different values of N, well.

So, N is in some sense a measure of co operativity, anyway I will come to that later. So, that is the idea. So, we have stat mech wise we have done very simple things writing down this

partition function and calculating the probability should be very easy, but at least from there we have managed to get the some sort of input, sorry output which tells me how this probability of binding is going to depend on the concentration of ligands, right. And this is, yes?

Student: (Refer Time: 36:09).

Student: (Refer Time: 36:10).

It can only go to delay of course,

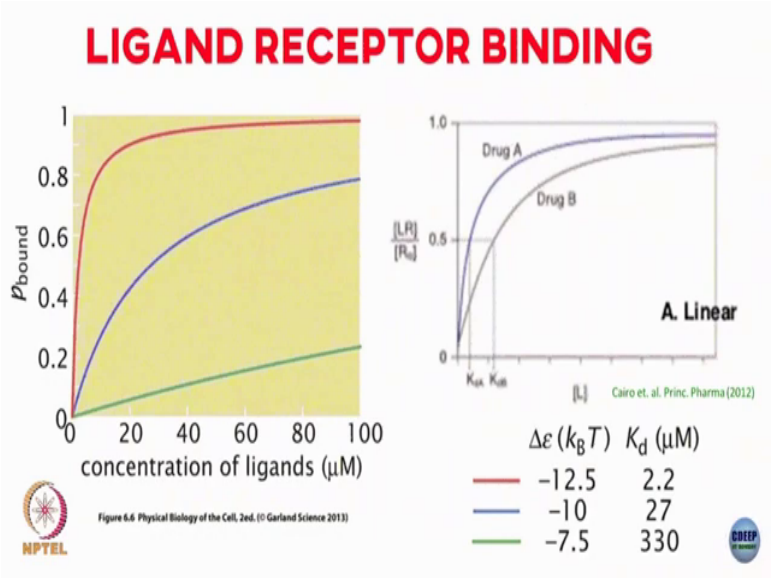
Student: (Refer Time: 36:13).

I mean, but N has is going up to infinity. Remember, this is a thermodynamic limit.

Student: (Refer Time: 36:22).

Yes, but this well to a certain extent that is true, that is true. Let us say whatever it will ideally depend. So, if you look at these sort of P bound plots exactly what we were plotting, right.

(Refer Slide Time: 37:18)



So, this is exactly what we said this is how the P bounds goes as a function of the concentration of ligands, these are for different values of delta E the higher the delta E. So, remember delta E, the sign will always be negative, because epsilon b will be smaller than epsilon solution, but as long as the magnet. So, if the magnitude of this delta is larger. So, green, blue then red and that is how the curves will rise up, ok. And in fact, this is indeed of in many binding experiments you can find curves like this.

So, for example, this was a drug trial where you were trying to ascertain the efficacy of two different drugs binding to a certain receptor molecule; this is in some pharmacological journal. And you will see that these drugs this is the binding probability the number of these complexes that you form LR, L is the ligand, R is the receptor as. So, this is the this corresponds to the probability of binding as a function of this ligand concentration. And you will see that the

curves look exactly as would be predicted by this in fact, what you can read off to a certain extent.

If you assume that this sort of approach is correct from these curves you can read off what would be the value of this delta E, how much would the energy we make lowered by the binding of drug A to the receptor molecule versus drug B to the receptor molecule, ok.

So, although this is a very simplistic approach, we have made many many approximations even. So, the overall the nature of these nature of these plots that you get from this simple calculation will often correspond to this binding, binding experiments for different ligands and different receptors, ok.

Of course, what you will be getting out if you were to get out a delta E of these experiments you will not really get the exact delta E at the molecular level, but some affective delta E which is going to be represented within the approximations of this sort of an approach, right.