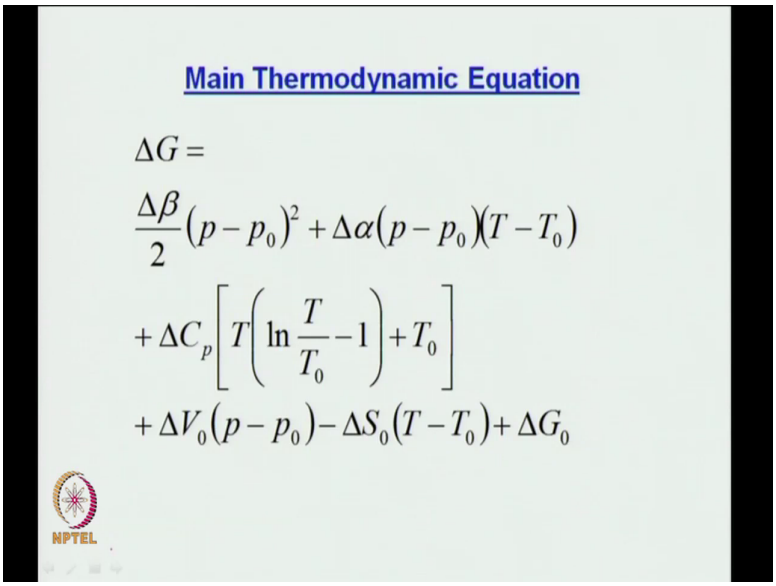
Bio-Physical Chemistry
Dr. Pramit Chowdhury
Department of Chemistry
Indian Institute of Technology, Delhi

Lecture - 19
Pressure Induced Denaturation (The P-T Diagram)

So, we will you know start off with where we left off from last class. It was the Pressure Induced Denaturation we were talking about. And we try to derive this expression for free energy change right, which has both the volume effects and temperature right, along with the respective thermodynamic variables.

(Refer Slide Time: 00:49)



So, this is the main thermodynamic equation that I wrote down. And just check one thing in the last class whether I had written delta G naught; hopefully I have not, I had not missed it in the last class. So, anyway this is total equation right.

Student: Sir, sir in this derivation.

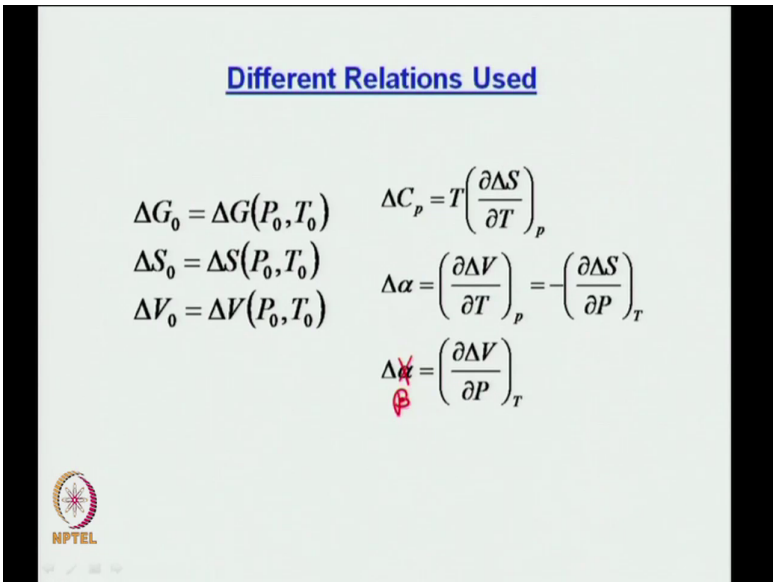
Right.

Student: Should not there be a 2 delta alpha there.

Right. So, what you do is, you take that to be delta alpha right; you take that to be delta alpha. So, that is I mean 2 delta alpha means, you just doing it the magnitude is 2 times more, right. But you do not do it in terms of in case of delta beta, because you will see that delta beta does not come twice right; this is we should actually getting from the integral. So, it remains as delta beta over 2. Yes you are absolutely right; so that means, you have done it ok, it was the good news for me.

So, again delta G depends upon you will see that delta beta, delta alpha, delta C p. We have delta V naught, delta S naught which are the constants, right.

(Refer Slide Time: 01:50)



So, what are the different relations we used in this case? So, the first one or the first column are the respective to delta G naughts delta, S naughts, delta V naught. So, this is a constant; at a respective pressure and temperature it would define as P naught and T naught right, it is for all the three. And then we also had delta C p is equal to this T of delta S over delta T at constant p, delta alpha is delta del V over delta T at constant p.

And which by Maxwell's relation gives you that respective relation with entropy; as a function of pressure at constant temperature and then we already have this delta alpha. So, this actually should be not delta alpha; this should be what?

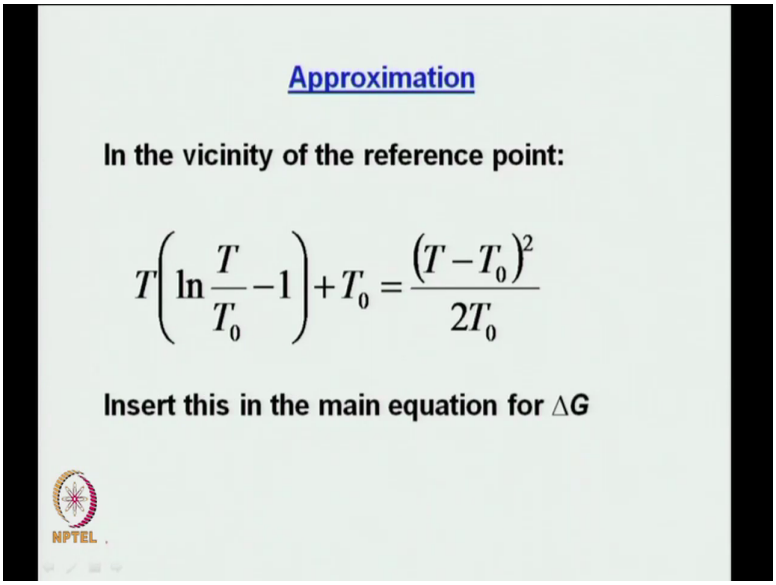
Student: (Refer Time: 02:36).

It should be delta beta.

Student: Beta.

Right. So, please make that correction, it should be delta beta; because delta alpha is already defined there, ok.

(Refer Slide Time: 02:46)



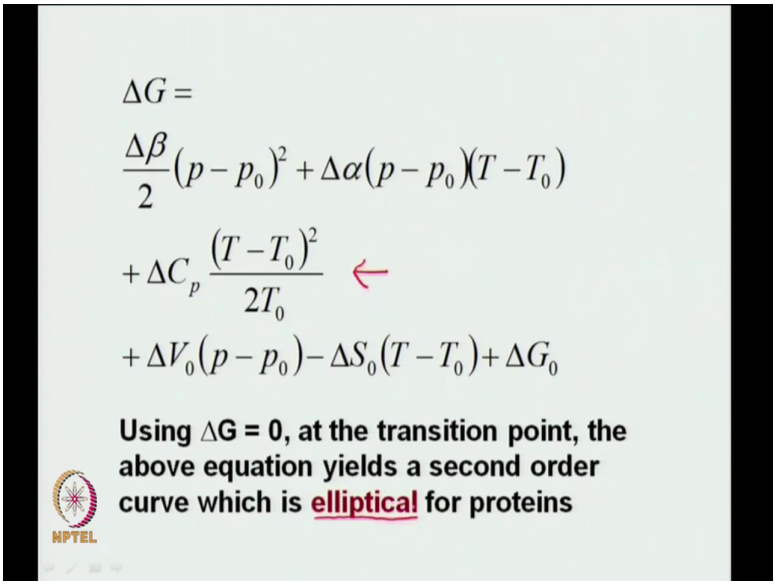
So, now, what we do is, having taken that delta G expression, we do an approximation; the approximation is in the vicinity of the reference point. We say, now this approximation we have seen before, we do a series expansion of natural log, right.

And then we can write this T of within brackets l n T over T naught minus 1 plus T naught is equal to T minus T 0 whole squared over 2 T naught or T 0, this is a reference temperature.

Student: (Refer Time: 03:13).

You insert this in the main equation for delta G and this is what you get?

(Refer Slide Time: 03:18)



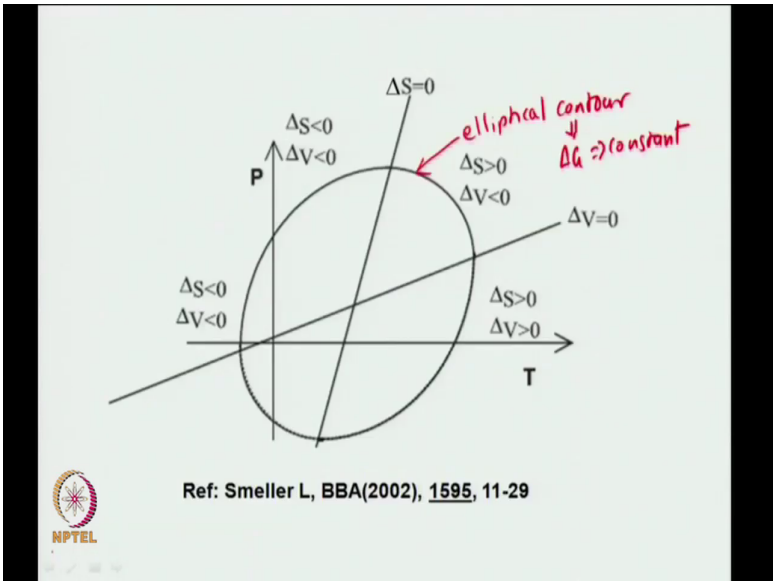
So; that means wherever you have this natural log term, the term with delta C p? So, this is you see this is where you have inserted the approximation right and this is your new equation, ok. Now what is the essence of this equation and why did we have to take this approximation into consideration? This is why.

If you would be using delta G equal to 0; you know where do you have delta G is equal to 0, at a transition temperature, right, at a transition point. If you would be having delta G is equal to 0; then at the transition point, the above equation yields a second order curve which is elliptical for proteins right and this is the what, this is a keyword, it is elliptical.

Student: (Refer Time: 04:04).

Which is elliptical for proteins right; you just put in delta G is equal to 0 and you will get something a relation between P and T, which is an elliptical relationship. And this is how the curve looks and this is how the curve looks.

(Refer Slide Time: 04:18)



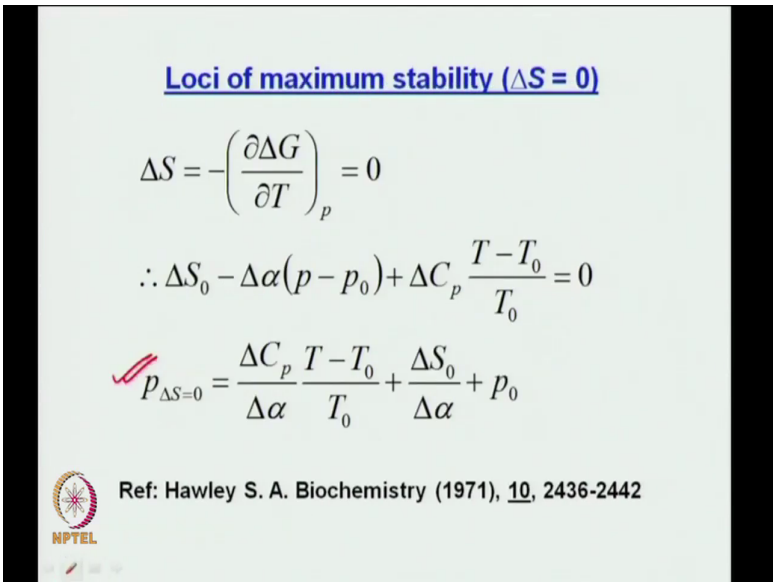
So, look at this, do not worry about these two straight lines; but look at this, this contour, this elliptical contour. So, this is the elliptical contour or contour, ok. Now, what does this

elliptical contour mean? This elliptical contour means, I will come back to it later again; but this elliptical contour what it means is, along the surface or along this perimeter of this ellipse, what do you have is, your delta G is constant.

Along this elliptical contour your delta G is constant, right. Now what do you mean by a constant free energy? Difference right, it is a constant free energy difference; that means, G of D minus G of N along all the points on the perimeter of the ellipse is the same. Now when can it be the same? Remember what is delta G equal to? Minus R T natural log K equilibrium; that means, no matter where you are your K equilibrium is the same, because delta G is the same, right. If your K equilibrium is the same, what is K equilibrium equal to? I will come back to this second latter slide; K equilibrium is equal to f of d over 1 minus f of d, do you remember.

That means this elliptical contour your drawing for a specific fraction of denaturant or denature state or native state. That means, your f of d is constant, the moment F of d is constant your K equilibrium is constant and then your delta G is constant. So, this is what we mean by the contour of this ellipse being or representing something where delta G does not change is constant over the whole parameter, ok.

(Refer Slide Time: 06:19)



Now what about these lines; delta S and delta V? How do we get that? So, this we refer to as Loci of maximum stability; you know what loci is right, it is a locus of all the points which define, which have this delta S is equal to 0.

Now, see what was delta S is equal to 0? Remember when we were talking about this hydrophobicity or hydrophobic effect we were saying; that the protein would be more stable at what temperature? T of S. And now what did we have at T of S? Delta S is equal to 0, that was the point of maximum stability, right. See if delta S is equal to 0; then I can obviously, write that minus del of del G over del T at constant pressure is equal to 0, ok.

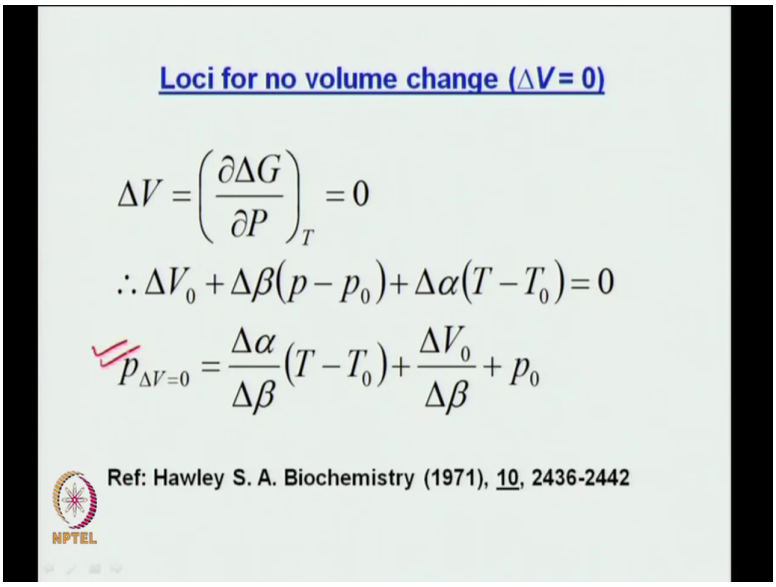
Now I already have a relational delta G as a function of pressure and temperature, we have already derived that. So, if you do that; if you do that. So, this is what you write down, you do the differentiation and this is what you get? You can see, this is the main equation; that

means, the pressure at any given point where delta S is equal to 0 has this relationship with temperature and delta alpha, ok. So, this line P where delta S is equal to 0 is essentially related to p naught, which p naught essentially is a intercept by the corresponding delta C p, delta alpha, delta S naught; and obviously it is a function of temperature T.

Because if you go back to the previous diagram, if you go back to the previous diagram or the previous figure; what do you have on the y axis? Pressure. What do you have on the x axis? Temperature; that means this ellipse is essentially what a pressure temperature diagram. It is essentially a pressure temperature diagram of your protein ok; here we are talking about proteins only. Now what does this say we will look at it later without going into too much of details; I just want to make sure that you look at it or at least know what it is. So, one was loci of maximum stability delta S; but see if you remember the delta G expression, I not only had delta S, I also had something which was delta V.

Why? Because along with temperature; along with temperature when you are changing pressure what happens is, you have a change in volume. So, you might also be getting a line where delta V is equal to 0 right; you know there might be a case where delta V is equal to 0.

(Refer Slide Time: 08:50)

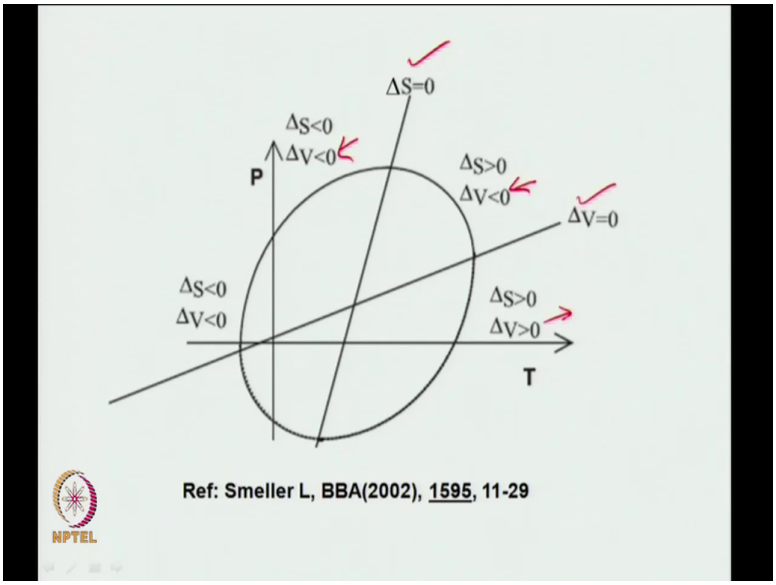


So, what you do is, to find that, to find that for no volume change that is delta V is equal to 0; we know that delta V is equal to del of delta G over del P at constant T, again we know what delta G is. So, we have we again this differentiate and you can see that, this is your loci of points where delta V is equal to 0.

Here again your relationship P with T, p naught is your intercept; but now you do not have delta C p here, what do you have? You have delta alpha, delta beta and the corresponding delta V naught; this make sense, because I am talking about delta V being equal to 0, I am not talking about delta S is equal to 0, ok. So, this is a easy to derive; you have the delta G, you just differentiate with respect to P at constant temperature for delta V and you will arrive at this expression, no problem, ok. So, now, hopefully realizing the significance of this, we look

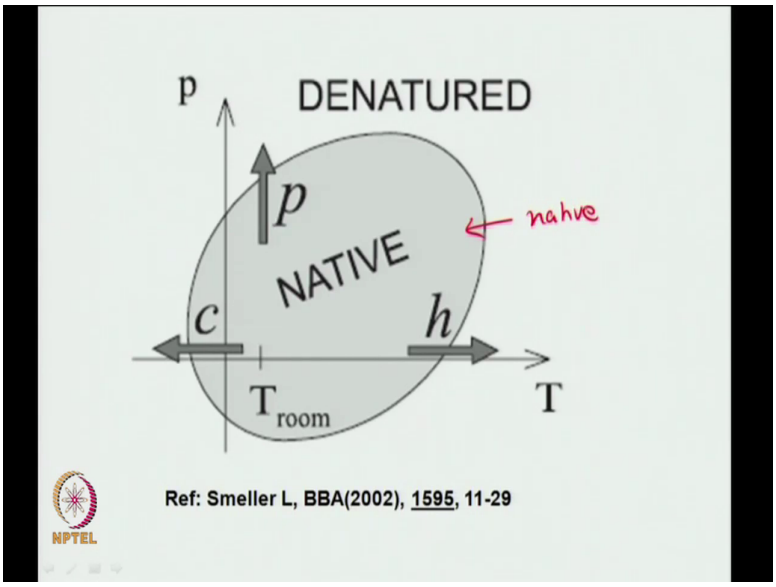
at this curve again. So, you can see these two lines; one is these two important lines delta V is equal to 0, delta S is equal to 0.

(Refer Slide Time: 09:46)



And look at what happens. On the left of delta S equal to 0 what do you have? You have delta S is less than 0. Now on the top of delta V is equal to 0, see this this is the delta V is equal to 0 line. And on top what is delta V? It is always less than 0; when you are on the other side of delta V, you have delta V greater than 0, ok. And this is just how you read the elliptical diagram; this is essential elliptical pressure temperature plot of a certain protein. Now what does this help us doing, do in. So, you look at the next plot.

(Refer Slide Time: 10:30)



Whatever region you had within the perimeter of the ellipse is the region where your protein is more stable; that means, the shaded region, you can look at the shaded region, you can look at the shaded region and it said there is a native region, right. This is a shaded region, this is within the ellipse; this is the phase or this is the phase space or the pressure temperature space, where your protein is stable, that means it is in the native state. The moment you move out of this elliptical phase region or the elliptical phase diagram or this elliptical region, your protein starts denaturing, ok.

See initially you had seen something like this where not exactly this, but you had seen a protein stability curve; do you remember that delta G was going like this, right. So, that was the stability curve, see that was going like this right; but there we did not take into account any pressure volume effects. But now we have along with that clumped into pressure and

volume effects and it gives us under certain conditions a an ellipse. And typically for proteins, this diagram is elliptical ok; for other molecules, other biomolecules it might be different, ok.

So, and this is almost like your, you know equal to I mean very close to a complete pressure temperature plot of an ellipse in terms of stability. And what you can see is from here, this what is C refer to, C refers to?

Student: (Refer Time: 12:02).

Cold denaturation, h refers to?

Student: (Refer Time: 12:05).

Heat denaturation and P refers to.

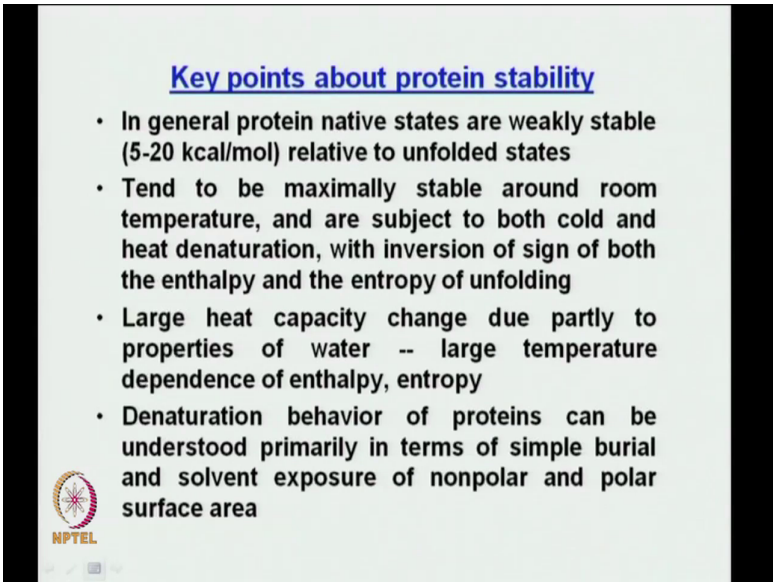
Student: Pressure.

Pressure denaturation right; so that means, in one plot you exactly know how much of pressure you have to put in at a given temperature to denature your protein or vice versa at a given temperature what or a given pressure what temperature would you know denature of the protein. So, that is why this pressure temperature plot is so important.

And remember when we are talking about denaturation of proteins; we are not talking about very low pressures, we are talking about pretty high pressures. What was it? It was all like close to about 0.5 to 1 Giga Pascal that is the range. So, that is pretty high, right. So, I have taken it from this references as you see I have jotted down at the bottom of the slide; but just keep this in mind; just keep this in mind, if you would ever come across such a diagram, you know at least you would know what it refers to or what it means, ok.

So, you know this is pretty much what I wanted you to know about your denaturation, I will not do p h as such; because I have already spent a lot of time on these different denaturation's. But before I move on to one of the main topics, which is you know actually how proteins fold the; you know that is we are talking about proteins right. I think now it is the time is ripe enough for us to go into the that main question, how proteins fold and what exactly is your protein folding problem. But before that let us look at small thing real quick; one is let us me go through this key points about protein stability.

(Refer Slide Time: 13:38)



The first and very important point is in general protein native states are weakly stable; only 5 to 20 kilocalories per mole more stable relative to the unfolded states. Appreciate this fact; appreciate this fact; see proteins are not rock solid, that means proteins are dynamic in nature, they will have conformational fluctuations. So, the point is if your stability of the protein is so

high; if suppose your stability is like say you know 100 kilo calories per mole, say 200 kilo calories per mole what will happen? If your stability is so high, it will become.

Student: Solid.

Solid. So, rigid right; that means, it will even if it has to make a small fluctuation, it would think about it at least a thousand times. Why? Because it does not have that energy; it needs a huge energy to make a certain conformation fluctuation, right.

So, that is why, but you know proteins are dynamic, they have to perform many functions; so they cannot be that stable. So, this is you look at the beauty of nature or the beauty of evolution; this is how they made the proteins. On one side you would not want your protein to denature very easily or get bad very easily, for that you would need a pretty high delta G; that means pretty high stability with relative to the unfolded state.

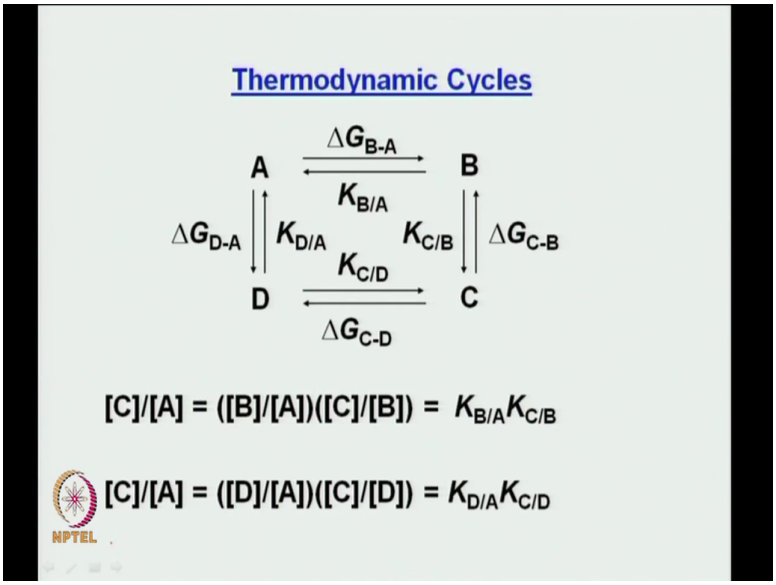
But on the other hand, you would also need a protein which is super stable; because that has to perform functions right, that has to be dynamic in nature, that cannot be rock solid. And we are talking about normal proteins, we are not talking about those proteins which are found in like high temperature organisms; there are proteins which like you know are exposed to very severe environmental conditions, but they still very stable, those are different proteins, right. Just normal proteins they would like to be stable and flexible to certain extent, not rock solid, ok.

So, that is the idea of the first point. Then the next is, this we have seen they tend be maximally stable around room temperature and they are subject to both cold and heat denaturation, with the respective signs getting inverted; we have not talked about cold denaturation that much in detail. Then one of the issues we have repeatedly talked about is the large heat capacity change which is partly due to the properties of water large and large temperature dependence of enthalpy and entropy; remember the slopes of delta h and T, delta S the slopes that were very steep coming from delta C p. And the denaturation behavior of

most proteins can be understood primarily in terms of exposure of buried hydrophobic groups.

Essentially what your accessible service area; that means we need denature, you make your hydrophobic groups more accessible to solvent and that brings about or has a huge contribution to your protein denaturation profile. So, this kind of sums up what you know protein denaturation or protein stability is all about.

(Refer Slide Time: 16:40)



So, this is the small thing I just wanted to point out before I move over to the protein folding problem. See let us quickly look at a thermodynamic cycle. Now try to recall where we look at a thermodynamic cycle; one was very early in the course remember where we are talking about a double mutant cycle right, we were do double mutations right that was a one cycle.

Then very recently we also look at a cycle right, where was that thermodynamic cycle do you remember? Hopefully you will remember by the time I am done with the slide. So, let us look at this. So, I have four states right and what I am saying is all these states are in equilibrium with each other; so that means, A is in equilibrium with B with corresponding free energy change delta G B A and the equilibrium constant K B by A right, and then B is in equilibrium with C with the corresponding free energy change same for C and D and same for D and A.

Now, have you been able to recall where you saw it.

Student: (Refer Slide Time: 17:45).

Ah?Student: (Refer Slide Time: 17:47)

Yeah We were doing this at transfer free energy remember?

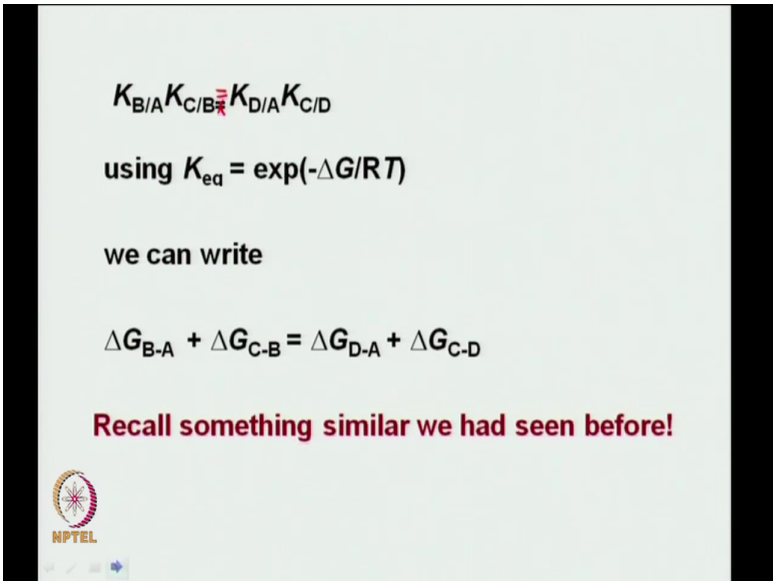
Student: (Refer Time: 17:51).

For this denaturation; well where g g t r, g t r comma i we are doing this transfer free energy that is where we saw it, ok. So, there we had we had written down a certain delta G expression, I just want to make sure that you know where it comes from. So, C if I have it like this what I can write is; since everything is in equilibrium, so concentration of C over concentration of A can be written as concentration of B over concentration of A times concentration of C over concentration of B, right.

Now, what is concentration of B over concentration of A? This is K B A is not it? And concentration of C over concentration of B is K C B; K being the equilibrium constant. Now I can also write C by A like this in terms of D now; I can write it as concentration of D over concentration of A times concentration of C over concentration of D, which is K of D in equilibrium with A times K of C in equilibrium with D, ok. Now where does this help us in? See what we are trying to equate is, we are trying to equate concentration of C over concentration of A, we have two different expressions. Now because concentration of B of

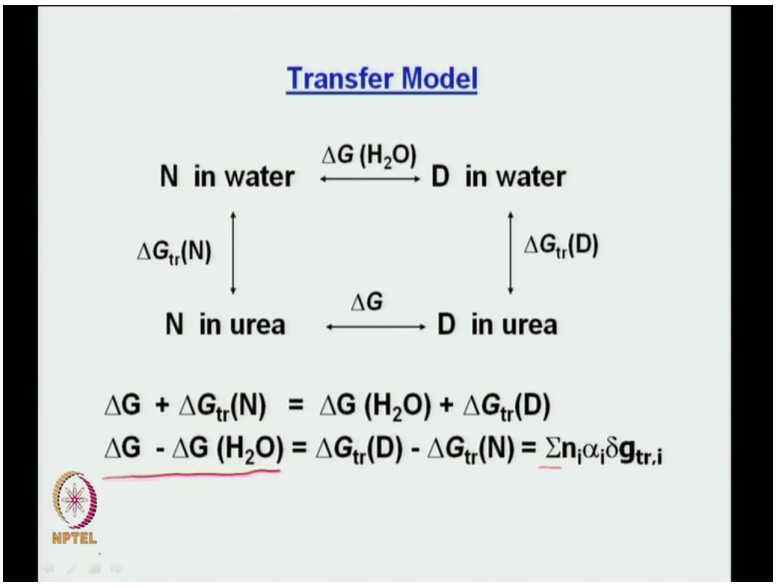
concentration of A is the same in both the expressions; that means, K B A times K C B should be equal to K D A times K C D, right. So, that is what you see in the next slide.

(Refer Slide Time: 19:15)



So, this equal to should be here, so this equal to should be here; so that means, K B A times K C B is equal to K D A times K C D. Now what does this give us? What does this give us? Remember what we try to or what we actually wrote down in one of the previous slides; I can use K equilibrium is equal to e to the power minus delta G over R T. If you do that, if you do that, would not this be the equation you would be getting; we can write delta G B A plus delta G C D is equal to delta G D A plus delta G C D, you know straight from here, right. We just do the simplification, does not matter how you do it and this what you are going to end up with, ok.

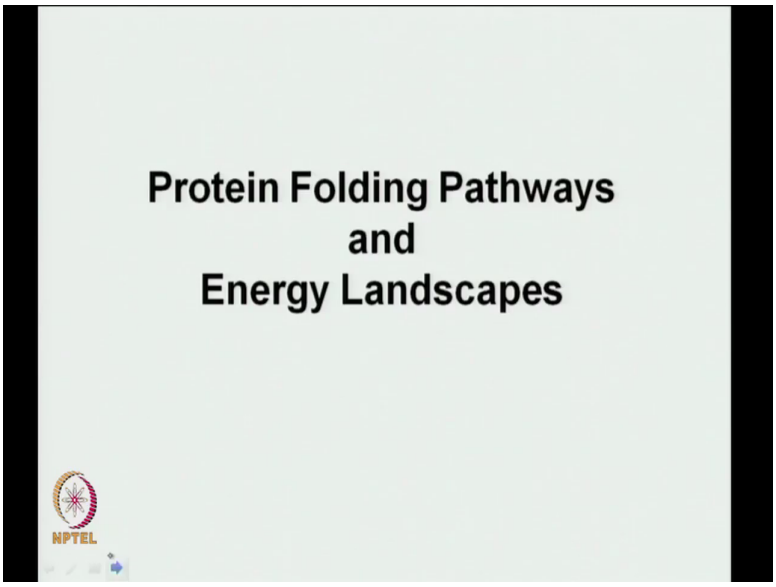
(Refer Slide Time: 20:10)



And that is what I said if you try to recall something similar we had seen before, this is what we had; we were looking at this transfer model right and we said for this delta G plus delta G transfer the native state is equal to delta G H 2 O plus delta G transfer of the denature state, that is exactly what you just derived from in the previous slide, ok. And you know that is what the final equation we had in that is around that slide, ok.

So, this is how a thermodynamics cycle is used and if you would be interested in deriving this p H expression, free energy change as a function of p H; then possibly you would be ending up using a similar thermodynamic cycle. Just look that up, I will not go through it; but just look that up see whether you can find something relevant to that, ok.

(Refer Slide Time: 21:02)



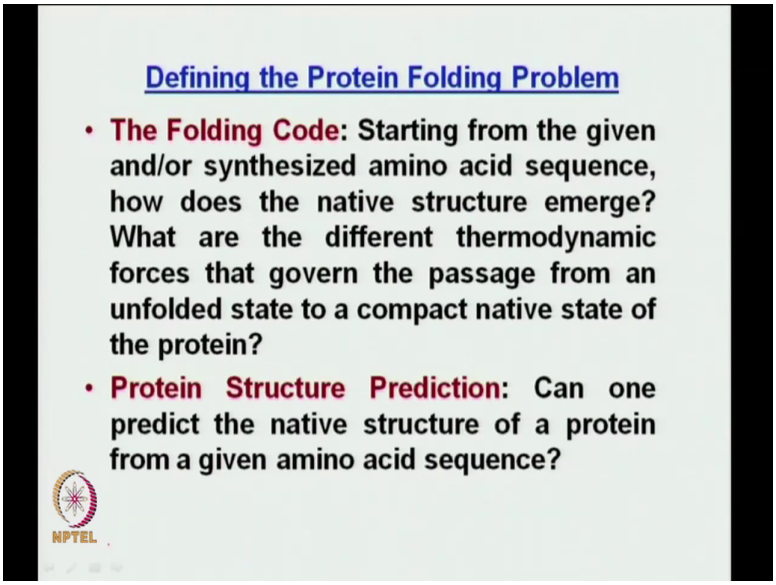
So, now let us come to the main topic, you know one of the most important topics someone can talk about or discuss about, if he or she is talking about proteins; protein folding pathways and energy landscapes. Now protein folding pathway is easy to at least you know realize; that means, if a protein when it synthesized in a ribosome, it is mostly without much of a structure.

And by the time it goes to the native state; that means, where it is in the compact functional form; I am talking about proteins with three dimensional structure, it must have crossed paths, certain paths, certain routes by which it came from the essentially non structured state to a state which has structure, right.

So, that is what we mean by protein folding pathways, what are the possible pathways; you know what do we have to keep in mind. And the other one is energy landscape; now this

energy landscape the meaning possibly might not be too evident right now, but I will I will let you know very soon, we will see what it means as we go along.

(Refer Slide Time: 22:14)



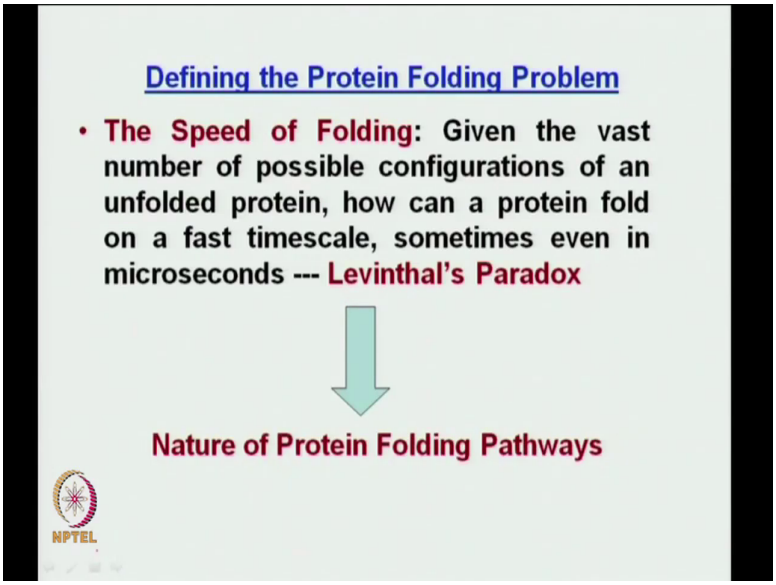
Now, let us define the protein folding problem. When we talking about a protein folding what are, you know how do we define it? You know what is exactly we are looking for? The first one is a folding code, now this we have already seen. What the folding code says is, starting from the given and or synthesized amino acid sequence, how does the native structure emerge?

Now when we say how does the native structure emerge, what we are saying is; what are the different thermodynamic forces that govern the passage or this passage from an unfolded state to a compact native state of the protein? Now the thermodynamic forces we have essentially looked at right; we have this electrostatics, hydrogen bonding, then one of the major ones

being the hydrophobic effect right, these are essentially the thermodynamic forces which allow the protein to go from the unfolded state to the native state. So, that is what the folding code is, right.

The second is protein structure prediction. Now this is also something we have referred to, but very briefly if you are given an amino acid sequence, can you predict the native structure of the protein? This is taking into account the fact that whatever final structure, the protein has is encoded where in the primary sequence; that means, the sequence of amino acids, that means if I give you a sequence just by looking at the sequence would you be able to say what is the final structure full alpha helix, full beta, mix alpha helix, alpha beta right like that.

(Refer Slide Time: 23:54)



And the third point is the speed of folding, right. See one was thermodynamics, one was prediction and the other one is kinetics; that means fast, how fast can it do that is also very

relevant question, right. Now given the vast number of possible, look at the statement; given the vast number of possible configurations of an unfolded protein, how can a protein fold on a fast time scale, sometimes even in microseconds this is referred to as Levinthal's Paradox. Now I will just pause for a moment and think what this statement is trying to tell you.

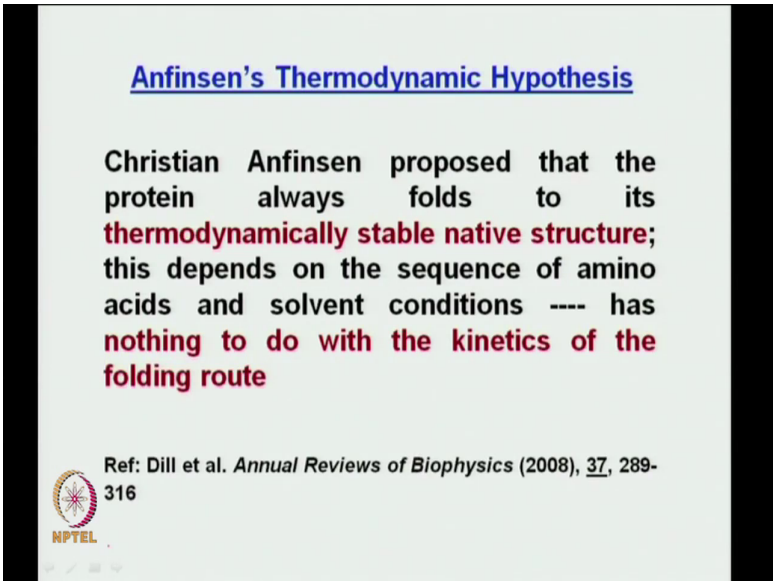
See when the protein is starting from its unfolded state to go to its folded state; what happens in the unfold state? In the unfolded state it is very flexible right; if it is very flexible, now think about there is nothing else just a protein right and the protein being flexible, it has no other information. So, what it tries to do is; it tries to sample all the possible conformations; because it is a it is very flexible right, being flexible you can easily go to many confirmations.

Now the point is if it tries to look at all the possible conformations it has, remember entropy is maximizing disorder. So, the protein would try to maximize the disorder, but trying to sample all the possible conformations. If we would try to do that, it would take a lot of time; because it would be having huge amounts of conformations. Think about all the dihedral angles the psi and the phi angles you have, right. However, even if a protein is doing that and as we discuss it would be taking a long time; but it has been seen that there are proteins which fold in very fast time skills, sometimes even in microseconds. So, generally if you talk about protein folding's, you can say some you know somewhere from milliseconds to seconds or you know few seconds. So, that is the time a protein takes to fold, but you would see in a later slide.

Now this time is rather quick; if you do a very cross estimation you would see that, this time is actually very quick, given the fact that the protein has a sample all the specific conformations it has or it might want to. What does this mean? Possibly it means is that, the protein might not be sampling all the confirmations; it might be having some push, some bias or some specific folding route that it might be following by bypassing all the configurations, it does not have to sample all the configurations or conformations, ok.

So, that is what the speed of folding is all about. So, then hence this defines your nature of the folding pathways; what exact path or what pathways do typically proteins follow, when they go to the final native state starting from a very highly disordered flexible unfolded state.

(Refer Slide Time: 27:07)



So, people have been thinking had been thinking about it for a long time. Now if you remember Anfinsen's hypothesis, this we did at the very starting of the class; it said that Anfinsen's proposed that the protein will always folds to its thermodynamically most stable native structure.

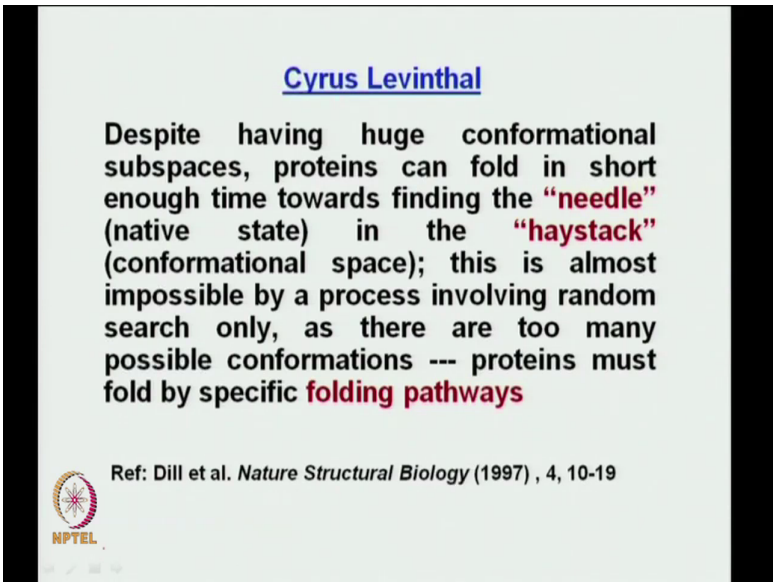
Now, there is a question about that. So, this depends upon the sequence of amino acids and solvent conditions. What do I mean by solvent conditions? Solvent conditions means that, if

you are taking a protein and if you are putting it in high concentration of denaturant and it would never go to the native state right; because you are always in a denaturing condition.

However, if you would remove the denaturant, then the protein would go to the native state. But what would define? It the native state would be defined by the thermodynamic principle; the one which has the minimum delta G, so that is Anfinsen's thermodynamic hypothesis.

So, what it means is that, all these delta G, delta H, delta S what are these? These are all state functions, right. So, the protein is not caring, what the pathways; what it only knows is the initial and the final state and being state functions, the pathways in between does not matter, so as long as it goes to the one which has the minimum free energy. So, that was Anfinsen's hypothesis, it did not talk about a pathway, right. So, that is why it says is a last line here; it has nothing to do with the kinetics of the folding route, how fast it does? You know there is no information out there.

(Refer Slide Time: 28:31)



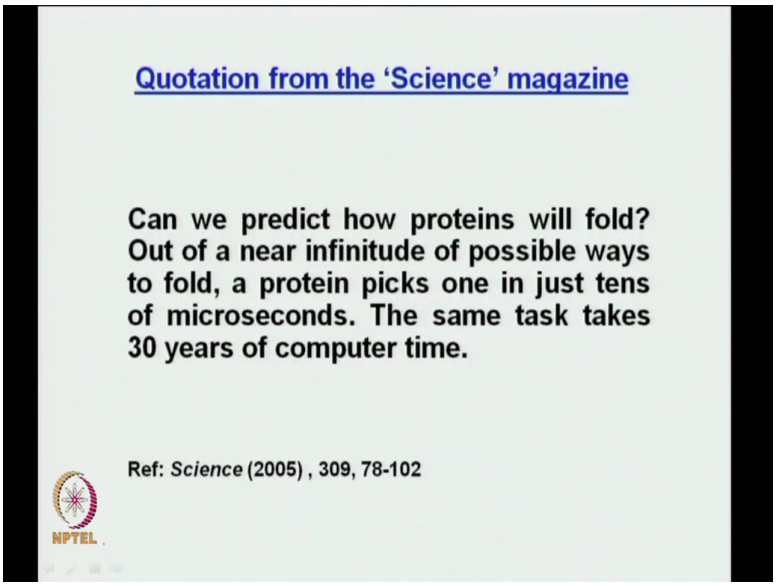
Now, Cyrus Levinthal again said this that, despite having huge conformational subspaces; I will tell you what a what I what we mean by a conformational subspace, proteins can fold in short enough time towards finding. Now this is a very important point, towards finding the needle; it is like a needle in a haystack problem, towards finding the needle in the haystack which is the conformational space. So, what do we mean by this? You think about the conformational space, the conformational space would be huge; because you have so many conformations the unfolded state has to go through, but by the native state has very little conformational entropy.

So that means, it would be say just one point, so it still does that; that means, among all the conformations possible it still just finds one global free energy minimum point, but that to on a pretty fast time scale. And when I talk when I am talking about time scale, I am talking about a biological time scale right, I am talking about a biological time scale. Now then it

says that it is almost impossible, a this is almost impossible by a process involving a random search only; because if we have to make this random search without any other information, then it will just take a lot of time.

So, it does it is impossible it is almost impossible by a process involving random search only, as there are too many possible conformations or configurations. So, proteins must be folding by specific folding pathways. Now this is what we mean by a folding pathway; that means, the protein must be following a specific folding pathway by bypassing this random search model of trying to search all the configurational subspace, ok.

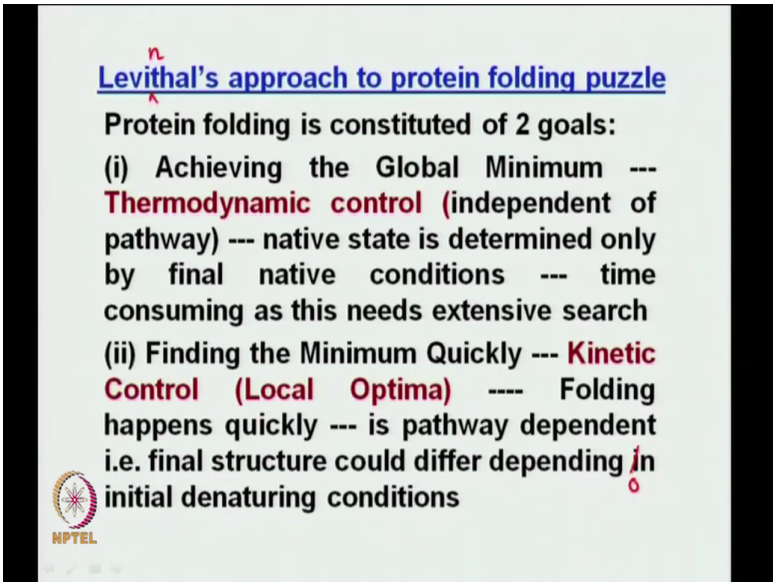
(Refer Slide Time: 30:21)



Now, this is a quotation from the science magazine, the year is given. So, this you know this issue of this particular paper was about the different problems in the world that need to be solved.

And one of the problems was this, can we predict how proteins will fold and out of a near infinitude of possible ways to fold are protein picks one in just; that means, out of the many many configurations a protein just picks one configuration say the native state and that too in just tens or microseconds ok, tens of microseconds. Now the same task if you would be asking a computer to do, it will take 30 years; again by this random search it is huge time. Now why would it take it, why would it take 30 years; and you will appreciate when we go to the later slide, now just hold on.

(Refer Slide Time: 31:22)



So, Levinthal, so there should be a n here, it is Levinthal. So, Levinthal's approach was two fold and this is how we are trying to solve this folding problem. The first is achieving the global minimum, we know that the native state is the most stable one right; so that means, the first one is a thermodynamic control. It does not matter what pathway the protein follows or

only matters is the initial state and the final state because these are state functions; as long as they as long as they reach the final native state, I am fine, I am happy.

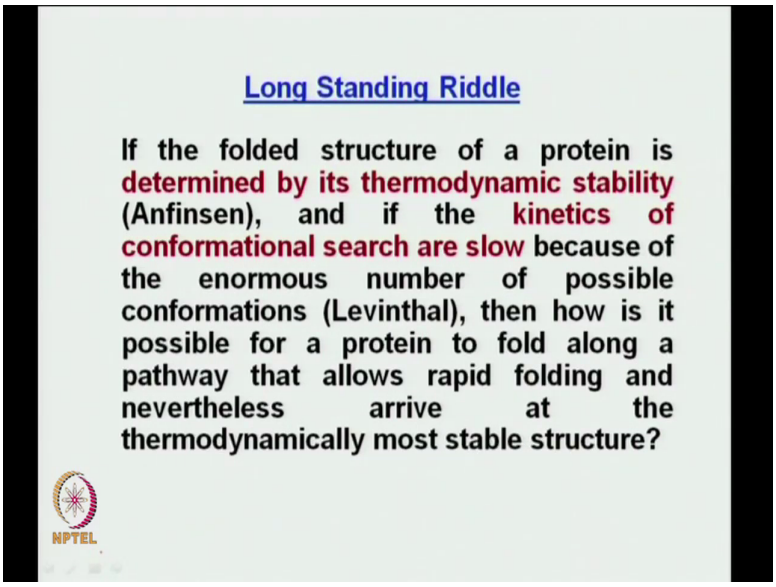
Now, if this is thermodynamic control, then in chemistry you already know along with thermodynamic control there is something known as kinetic control right, which you know about. Now if you talk about the kinetic control now, which means finding the global minimum quickly. Now kinetic control means that, folding happens quickly; it is depend on pathway, but del G and all these things are state function does not depend upon pathway right, these are all state functions, it does not matter what pathway you take.

So, when you are talking about this number 2 the finding minimum quickly it says that, the folding happens quickly, it is pathway dependent and the final structure could differ depending in depending on initial denaturing conditions. So, it should be on, depending on initial denaturing conditions. Now what it means is this? Here it is not talking about a thermodynamic minimum; it is saying that based on a specific pathway, it can reach a certain minimum which is not a thermodynamic minimum, instead it can be a?

Local optimum, but not necessary the global minimum.

So, how do you reconcile these two? one is reaching the thermodynamic minimum and the other one is doing it at a very fast time scale by doing a random search. And Levinthal said that these are possibly two separate events, exclusive events they personally do not have a relation with each other. Now the question is, can you somehow relate these two by some measure or the other ok?

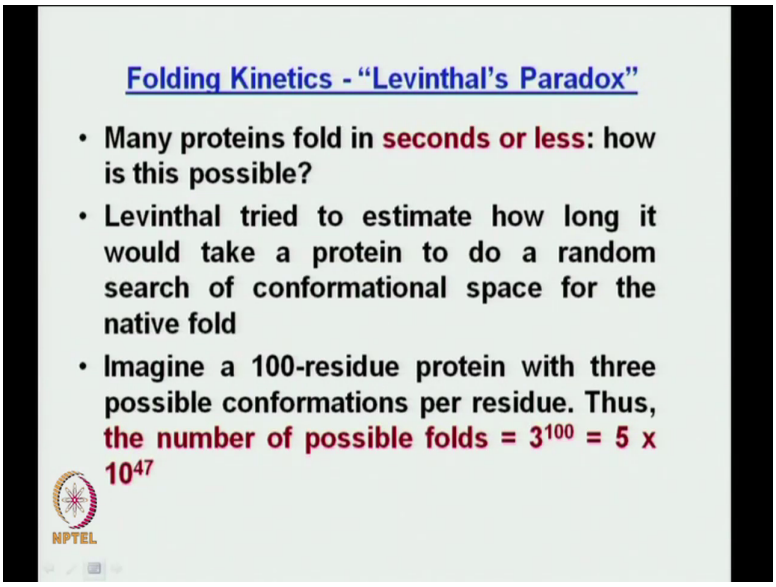
(Refer Slide Time: 33:40)



So, again this is another form, the long standing riddle. So, it says that if the folded structure of a protein is determined by its thermodynamic stability as told by Anfinsen and if the kinetics of conformational search are slow because of the enormous number of possible conformations. Now we are talking about the speed as Levinthal said, then how is it possible for a protein to fold along a pathway that allows rapid folding and nevertheless arrive at the thermodynamically most stable structure? This is precisely the problem right; you have thermodynamics and you have to do it on a short enough or fast enough time scale, this is essentially the problem that Levinthal faced.

And this was referred to as Levinthal's paradox, because this was a paradox right; he could not quite match these two seemingly separate things, ok.

(Refer Slide Time: 34:38)



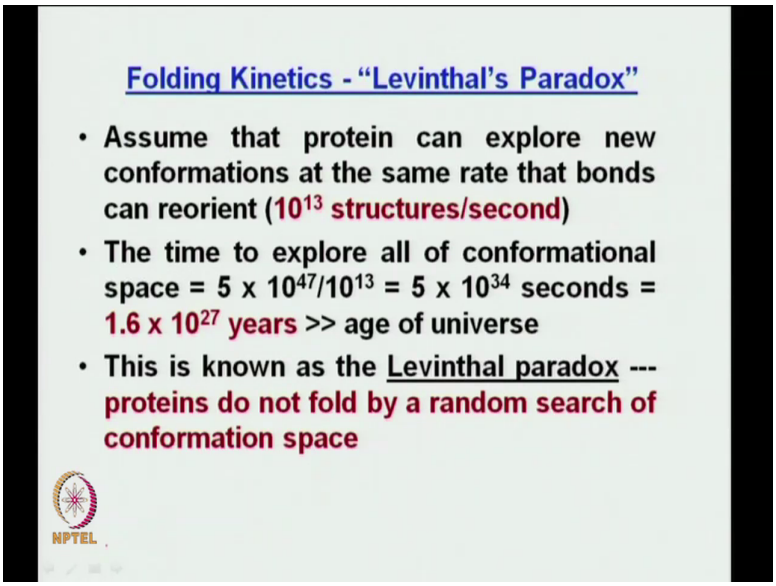
So, now let us look at some numbers, I hopefully will appreciate what we have been discussing this far. Many proteins folds in second, you know fold in seconds or less; how is this possible? So, Levinthal tried to get a rough estimate, ok. And how do you try to get the rough estimate? What he said is, let us do this; I take a 100 residue protein right and I say that let each residue have a minimum of three possible conformations and you can have more, but let us say it is a minimum of three conformations right, it is minimum we are taking.

Then the number of possible folds, remember each possible conformation constitutes a fold and you have 100 amino acids right and each conformation for each amino acid would be a different fold, is not it. That means, the number of possible folds would be what 3 raise to the power 100; because 100 amino acids and this is what you get 5 times 10 to the power 47. So, these are so, 5 times 10 to the power 47 is based on this estimate the number of conformations you have; remember this is your configurational subspace, the number of or

conformational subspace, the number of conformations that the protein will have to go through.

But this is a minimum estimate right, because I am just taking three conformations per amino acid, ok. Now what does the time come in?

(Refer Slide Time: 36:02)



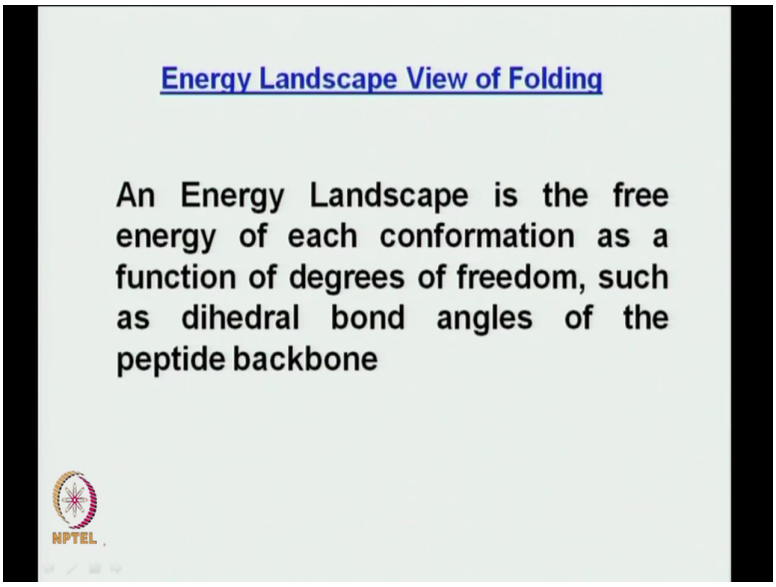
Now we assume, now this is also on the higher and we assume that each conformation can be explored at a rate which is about 10 to the power 13 structures per second; that means, you explore 10 to the power 13 structures in one second. That means, each structure exploration is done in what time? 10 to the power minus 13 seconds is not it; that means, each conformational change, here exploring in a time of what 10 to the power minus 13 seconds, right.

And that is how you explore 10 to the power 13 structures in one second. So, you know how many structures you have; 5 times 10 to the power 47, you know the time right 10 to the power 13 structures per second. So, now, you can easily do this math and see; the time to explore all of the conformational space is what? 5 times 10 to the power 34 seconds; see but remember this is a minimum estimate right, minimum.

If I am going to take much more confirmations go even higher right and this is what you are going to take 1.6 times 10 to the power 27 years which is almost the age of the universe. I mean if that had been the case, then none of us would have been here right; I would not be I have not been talking about bio physical chemistry neither had you been sitting and listening to what I am talking about, right. So, you see there has to be a balance somewhere and how does the protein do it, ok?

So, again this is known as the Levinthal's paradox proteins. So, the bottom line now is the main message, the proteins do not fold by a random search of conformational space; because if it tries to do that, it will just take up a lot of time and we know from practical experience that proteins do not do that, ok.

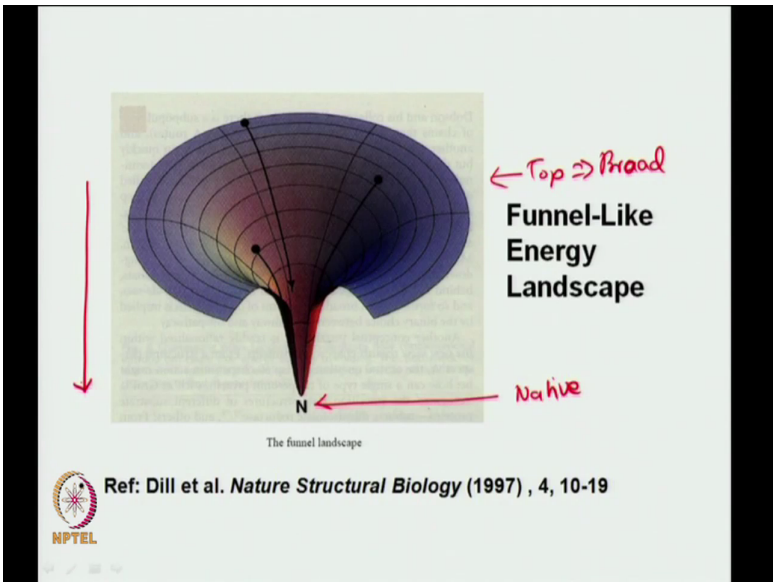
(Refer Slide Time: 37:56)



So, how can these facts be reconciled? So, then what people said, they talked about something know as energy landscape view of folding, right. I will very quickly I will show you a figure which talks about this energy landscape view. Now, what do you mean by an energy landscape?

So, an energy landscape this is how it is defined; is the free energy of each conformation as a the function of degrees of freedom, such as dihedral ball angles of the peptide backbone, ok. Because remember of, anyway when we are talking about a protein coming to the native state, it is starting from a very free flexible state; you would be having a whole range of dihedral angles and the energy landscape is trying to capture that.

(Refer Slide Time: 38:44)



So, this is what you looking at, it is a funnel like energy landscape and try to realize what it means.

You look at this, you look at the top of the funnel ok; that means, you look at this top of the funnel here, look at the top. The top is what? Very broad is not it? Then look at the bottom, what do you have here?

This is your native state. Now this is like a funnel; because remember in a funnel what do you have? In the funnel it is very broad at the top and then it tapers down and comes to the stem, right.

Now, obviously here we are not talking about the stem, but the point is that, at the top is very wide and then slowly it narrows down and comes to the native state, right. Now it makes sense, why does it make sense?

Remember what are we starting from? We are starting from the denatured state; denature state, what do the denature state have? High degree of flexibility; if the denature state is having high degree of flexibility and try to realize what your x coordinate is now, then what is the x coordinate? Your x coordinate is actually your measure of flexibility; measure of flexibility means what? Measure of entropy essentially. So, x coordinate is essentially entropy; so that means this is the breadth, this is the and this is the breadth of entropy or this is the wide distribution of flexible conformations you have in the unfolded state, because remember you are starting from here and then you are moving down.

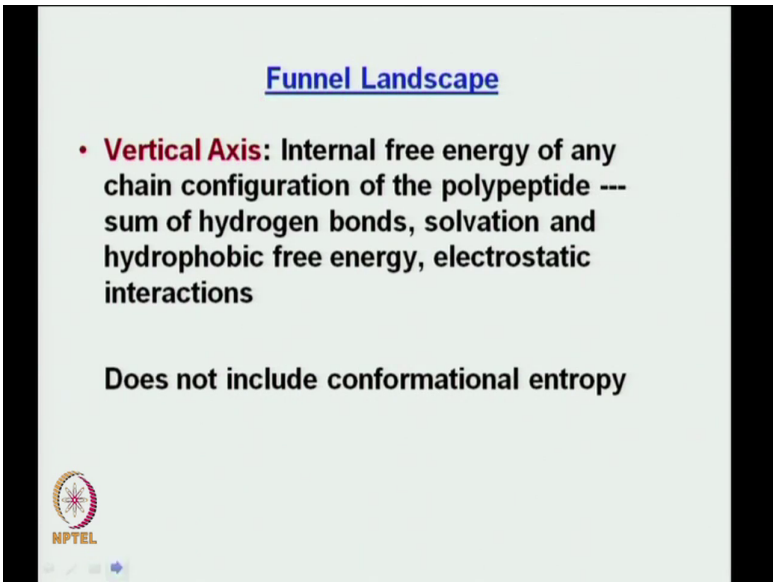
And on the y axis you have your

Student: Refer Time: 40:23).

Internal free energy, I tell you very soon again. So, again try to appreciate this funnel like landscape. Funnel like landscape means that, just looking at it means it has so many conformations to start with. Now as it starts folding what happens is, you reduce, reduce, reduce, reduce, reduce so number of confirmations right; and finally go to the native state which is like a unique configuration. And hence your entropy of the system has decreased so much and that is typically what happens in a funnel.

Now, the thing is what you can immediately ask is; why do I just consider funnel, why do I not just stick to one side and start from there ok? Just hold on again, I will tell you why we really need it to be funnel like.

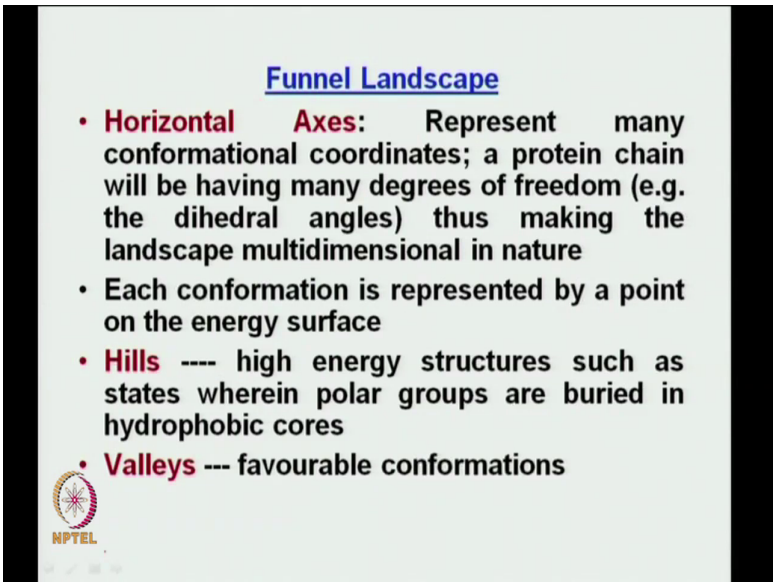
(Refer Slide Time: 41:12)



So, in this funnel landscape, what is my vertical axis? It is the internal free energy of any chain configuration of the polypeptide; that means, see when the protein is coming back from the coming from denature to the native state, what is happening? It is gaining interactions right; hydrogen bond, Van der Waals forces, salt bridges, dispersion forces, hydrophobic forces all these things. So, it is getting stabilized by those forces, it is mainly what enthalpic, right.

So, it is a sum of all these stabilizing energies right that is the vertical axis. But what it does not include? It does not include conformational entropy; it is your internal free energy divide of this entropy contribution, it is essentially your enthalpy contribution, ok.

(Refer Slide Time: 42:04)



Then, what is your horizontal axes? This horizontal axes, now I say I do not say axis, I do not say a x i s; I say axes. Now please try to get you know the meaning of this; axes means, multiple axes. Why do I say multiple axes?

First of all these horizontal axes represent many conformational coordinates; a protein chain will be having many degrees of freedom right, example the dihedral angles. The dihedral angles you have many angles right; thus making the landscape multi-dimensional in nature.

Now do you do you understand this point; when we are taking about the x axis now, you will not you are not talking about this one coordinate, you are not talking about one dihedral angle. You are talking about what? Each and every residue would be having what; many many each and every residue would be having their respective dihedral angle.

So, what you can do is, you can take one residue out of the 100 amino acids; you can say take the 75th residue and follow its dihedral angle as the function of the protein coming from the unfolded state to the folded state that is one coordinate. You can take another protein, you can follow another dihedral angle; I am sorry not on the protein, on the residue I mean and follow another dihedral angle and follow again. So, what will happen? Your total conformational space, so your total energy landscape is what; it is a summation of all the contributions of all the dihedral angles. And that is why it is not just one reaction coordinate or x coordinate, what is it; it is a multi-dimensional coordinate.

That means you have many many dimensions, that is why it is multi-dimensional in nature; because it just do not have one x coordinate, you have many many x coordinates clumped together in one energy landscape feature. Do you understand why it is multi-dimensional now; it is not one dimension, because you are talking about many different separate things right, many different residues, separate residues, separate dihedral angles, separate requirements all these things, ok. So, then essentially your horizontal axes is your entropic.

So, each confirmation now is represented by a point on the energy surface. So, now, that makes sense; if you have a plot, like you are having a pressure temperature plot, then one at one point what you will get, you will get the corresponding pressure and the corresponding temperature. Similarly the energy landscape each point would be giving you the corresponding conformation; say the phi angle, psi angle whatever, ok. The in the energy landscape there are two things that you might come across; one is hills, hills means this is the high energy structures hills are like barriers, high energy structures.

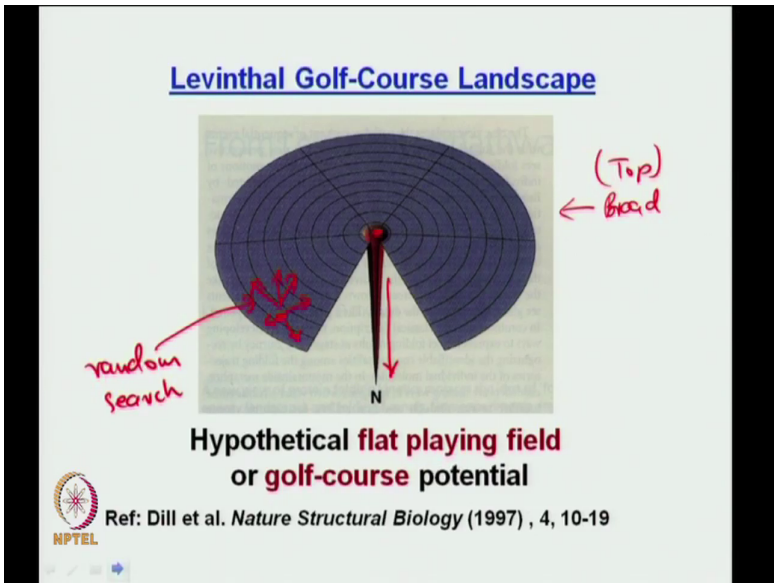
How can you have high energy structures? You can have high energy structures for structures which are not that stable; that means, suppose one of very you know easy example is see the proteins are folding randomly right in a sense, so they would not notice. So, in a certain conformation would they might happen is or it might happen is that, they might be having a host of polar residues pointing towards the hydrophobic core; but then that is not conformation is stable or that is not energetically stable right, because you cannot have polar

residues in hydrophobic core, see immediately the protein will open up. But obviously, the energy of that configuration is what low or high?

Student: High.

It would be high, so that is a hill. Then obviously, if have a hill, you will also be having valleys, right. So, valleys would be your favorable configurations, where you have all these interactions, the polar groups are not inside, but pointing outside, the hydrophobic groups pointing inside like this. So, this is essentially your energy landscape, ok.

(Refer Slide Time: 45:24)



Now how do I sketch it? Look at this golf course landscape. Now this is called Levinthal golf course landscape; I will actually end here after discussing with golf course landscape. Now

this is very important. Why is it called a golf course landscape? Look at this, if you look at the top; you look at this top, it is so broad, right.

The top is so broad again; but there is one problem, the problem is for the protein, for the protein to go from any point on this top surface. What is the top surface? The top surface is it is unfolded state; it is broad because, remember every point on the energy landscape is what, a specific conformation and that is why in the unfolded state you see a huge range of conformation that is why it is so broad, right. But the problem is see; if only the protein conformation is somewhere here, if only the protein conformation is somewhere here, then it can come directly to what the.

Student: (Refer Slide Time: 46:33).

Native state. What I mean is again, try to realize the importance; only the protein conformation suddenly corresponds to this coordinate or to this point, then the protein immediately sees that I have a downward slope towards my native state, then you can easily go to the native state. But here is the problem; the problem is what have we said, there is no guarantee that the protein would be in only one set of or in only one conformation, it would be having many many different conformations, thousands of conformations.

So, what about the protein say which start from here? See, what about the protein which starts from here? Either the protein is going to start from here, see the protein can go here right, the protein can go here, the protein can go here, the protein can go here the protein can go here; but there is no easy route. What do you mean by easy route? The protein if it is going to start from here at the extreme end, it would not know that the native state; to go to the native state I have to go to the center which is the golf part. Why? Because your what is your energy, your energy surface is flat.

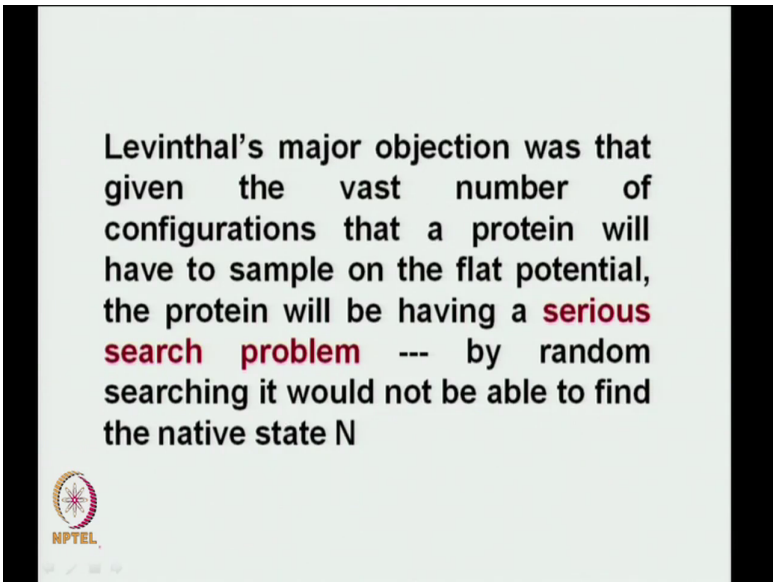
If your energy surface is flat, it is like a ball rolling like a golf ball rolling;, the ball will just roll irrespective of where it is; that means, there is no hill, it is a plain, right. So, there is no bias for the protein to go to any specific direction, it can go in many different directions; the moment the protein will go in many different directions, what. See what is this referring to?

This is referring to what, is not this your random search? This is a random search right, because the protein has no information to how, to go to the native state.

It would keep looking; it would keep looking for that specific golf parting hole or this hole in the middle which will take it to the native state. And to do that, see it has no force, it has no direction; why? Because each and every point, each and every point on the landscape has equal probability; because it is the flat landscape, is not it. Like if you let the ball roll here, if it is a flat, if you let the ball roll here; the ball can go this direction as I said, the ball can even come into this direction, nothing is stopping the ball from going in this direction, same for the protein, nothing is stopping the protein from trying to sample all the spaces.

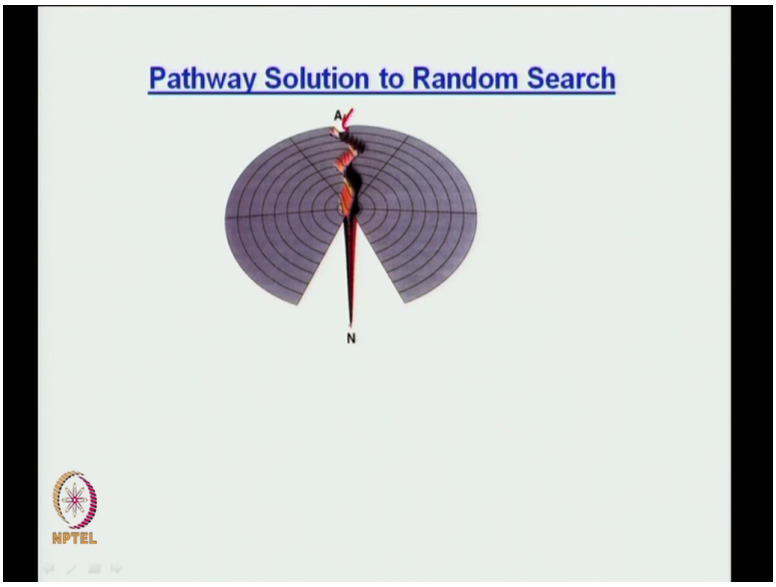
Again this is your random search. So, you can imagine now, how long will it take? It will take huge amount of time; then obviously, the point you can make is possibly this golf course landscape is not the proper one to look at, because if the protein would be following this, it would be following a random search model. And trying to follow the random search model what will it do? It would take up just huge, huge, huge amount of time. So, the protein definitely would not be doing that, ok. So, this as I say, it is a hypothetical flat playing field or golf course potential.

(Refer Slide Time: 49:24)



And what was Levinthal's major objection; he said that given the vast number of configurations that a protein will have to sample on the flat potential, the protein will be having a serious search problem, by random searching it would not be able to find the native state N. And even if it does finally finds the native state N; you will see it already has spent a huge amount of time, ok. Now this was one, now this is but I will stop; but I will just show you where we are going to do, where we are going to go next. You look at the next one.

(Refer Slide Time: 49:54)



You look at the pathway solution; pathway solution is what does this say? It says that, if the protein is going to start from state A. Look at state A, what this had? It kind of has a valley, a tunnel.

Student: (Refer Time: 50:07).

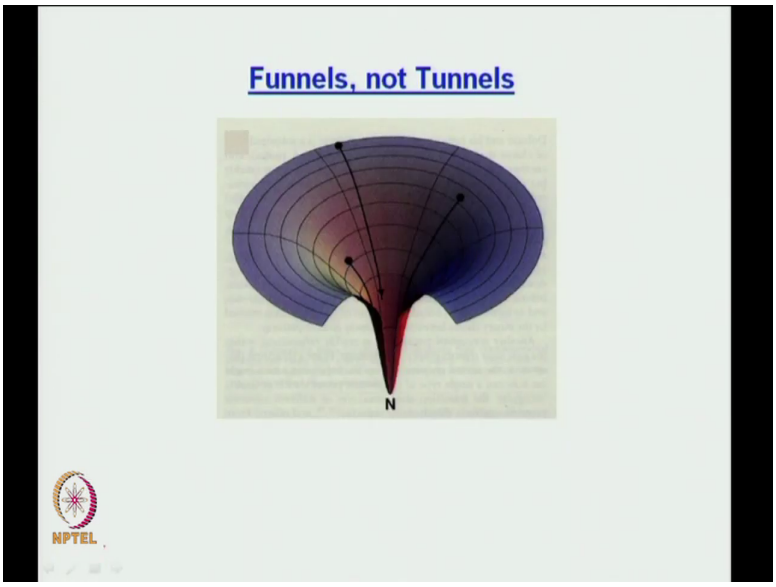
Say, if it starts from this state, it will come and go straight to N. So, these are path basis; that means, there is already a pathway made for the protein, that if it is there it will go to the native state.

But then again where is the problem? The problem is what this is saying is that, it is the protein is only in one denatured state which is A, and we know that cannot happen.

So, if the protein is somewhere on other portions of the landscape, again it will face the same random search problem, ok. So, this is called as specific pathway model, we will start from here again tomorrow.

Then think, now try to think what would an actual energy landscape, a feasible logical energy landscape look like? A feasible logical energy landscape I will just show you, a logical one, possibly would look something like this; that means it does not matter where the protein starts from, any given configuration.

(Refer Slide Time: 51:04)



So, this is your configuration space, any given configuration starting from any point on the lands flat lands on the; not flat on the landscape can go to the native state ok,

This is possibly a logical funnel or a logical energy landscape, right. So, we will start from here again tomorrow.