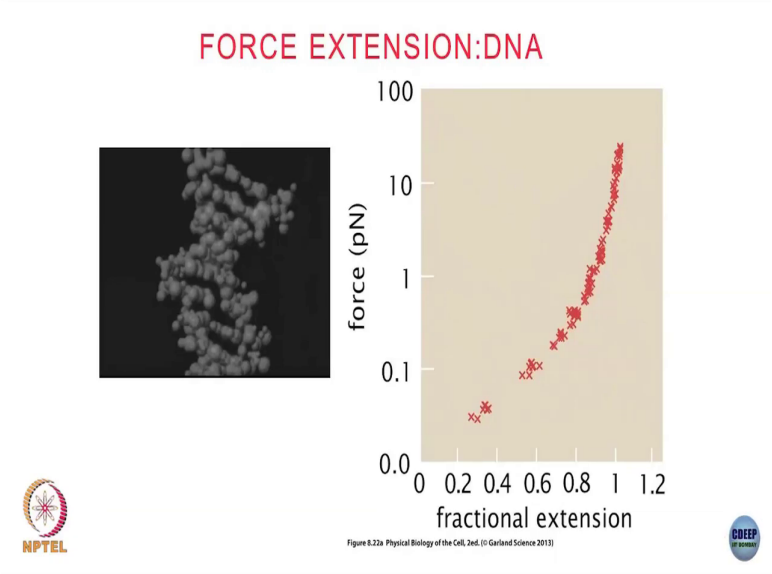
Physics of Biological Systems
Prof. Mithun Mitra
Department of Physics
Indian Institute of Technology, Bombay

Lecture – 36
Random Walk models for Proteins

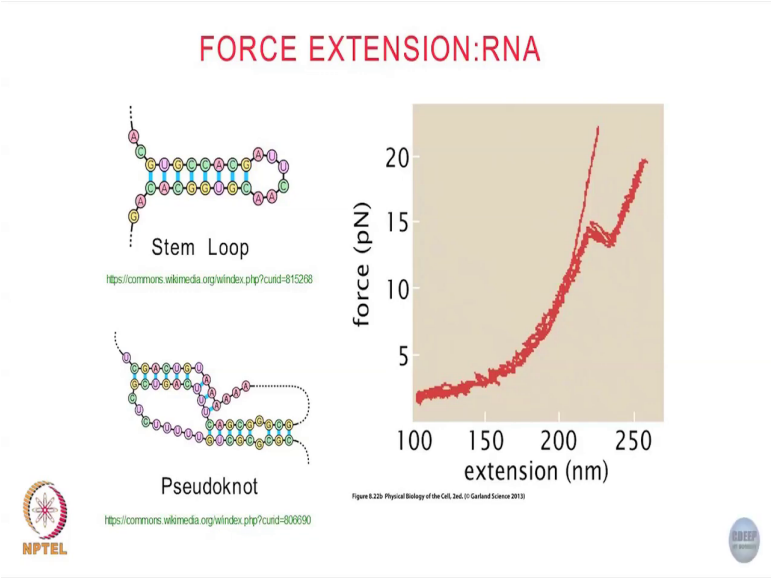
So this gives us a way to sort of fingerprint different polymers, different sequences, of DNA, but not only DNA you can do this to any polymer for example, and that is what I just want to show.

(Refer Slide Time: 00:35)



So, this was force extension of DNA, want to do a force extension for a protein ok.

(Refer Slide Time: 00:47)

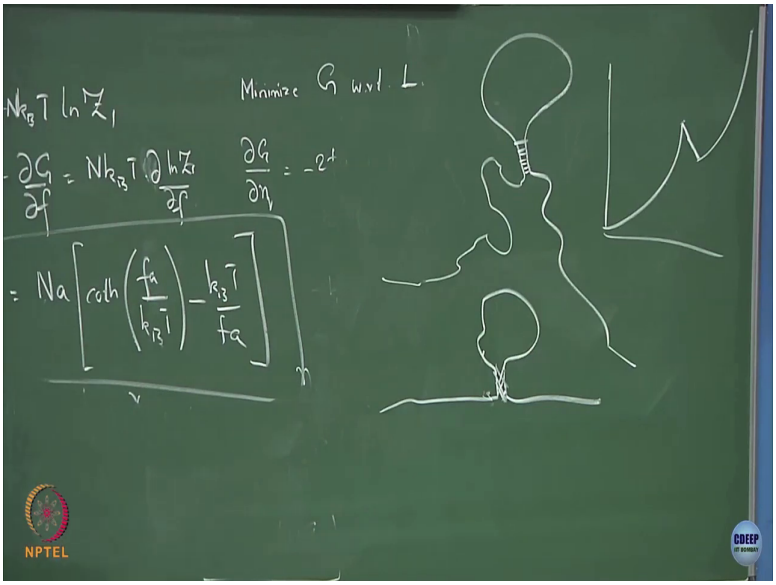


So, I will come to protein. So, first let us look at force extensions of RNAs. So, here is the typical sort of force extension curve, for a pulling in RNA molecule. And, you some instead of; so, its looks roughly similar like this DNA, but then there is something going on over here ok. If you saw a force extension curve like this what would you, could you guess anything about what the structure what sort of a structure of this polymer could lead to force extensions like this.

So, basically if I look at it what is happening is that I am pulling on it as I am, as I am pulling on it my polymer is extending, but it is getting sort of more and more difficult in order to extend which is why the curve goes like that. But, then after a certain point it becomes slightly easier to pull a certain length of the polymer right.

So, what that would correspond to is something like this, remember RNA is not a double stranded (Refer Time: 01:45) right. So, if you have complementary base space so, this was my original RNA strand, but if some segments have complementary base space you can exactly like in DNA you can form some interactions between these different segments of the loop.

(Refer Slide Time: 02:06)



So, if you were if you had something like this for example, I had a polymer and then I had a segment over here which had sort of form these base pairs right, and I am pulling the ends of this polymer. So, as I pull it this thing gets extended so, I get something like this maybe right. So, when it is in this sort of a conformation that is where this curve looks something like this, it is now gotten difficult to stretch.

But, now once I have forced let us say I, I pull with sufficient force that I break open these bonds that were there, then suddenly with very little force I can open up the polymers a little,

little when these bonds break. And so, I can open up this polymer with very little force which means I get a sort of kink here, then once this domain has opened up I go back to my original, my usual sort of force extension behavior ok.

So, typically therefore, in RNA you can have different sort of structures of course, you can have loops, you can have knots and so on; depending on the sequence of the RNA you can have multiple you can have multiple such things and so. So on but typically so, therefore, a force extension curve would look something like this.

Student: (Refer Time: 03:21).

No, so, this curve is for a specific RNA sequence.

Student: (Refer Time: 03:32).

How did it have?

Student: (Refer Time: 03:38).

So, let us say which extension point should I be looking at?

Student: (Refer Time: 03:50).

Student: (Refer Time: 03:52).

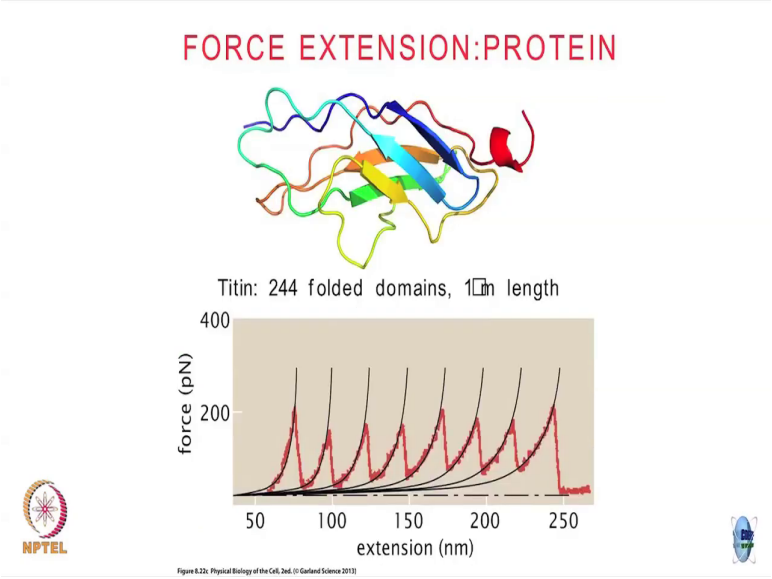
This thing?

Student: (Refer Time: 03:55).

That is just our guide for the eye, that if I were to do something like a DNA this would look something like this which is why this screen is very bad. These are actually points that is just a

line that has been drawn, so the points of the actual experimental data. So, pulling on RNA molecule would look something like this; on the other hand let us say I pull a protein now.

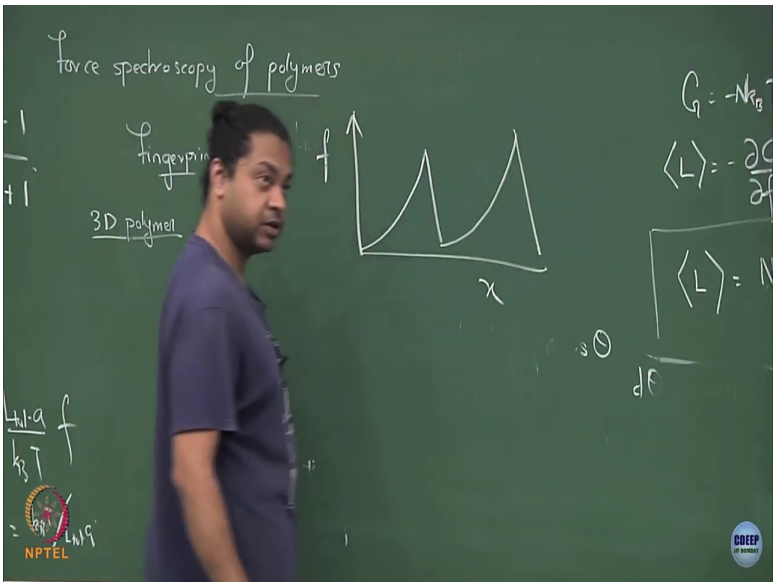
(Refer Slide Time: 04:21)



So, let us say I pull a protein, for example we were talking about this Titin protein which is one of the largest proteins, it is 244 folded domains, it is a multi domain protein. And I say that I want to pull this Titin protein. So, again I draw this force versus this extension sort of a curve. What would a curve for a protein like this look like? What would the force extension curve looks like?

Student: (Refer Time: 05:04).

(Refer Slide Time: 05:14).



A lot of steps and steps in the sense that you pull, once you have reached one of these domains, the moment you have managed to unbind that domain; you need now very little force to pull it little more and then you repeat this sort of a structure and indeed that is what you see. So, if you look at it this force extension relationship this is what the force extension relationship looks like.

And in fact, if you so, if you were to take a random protein and you were to pull it and you got some curve like this; a force spectroscopy so, a force extension signature like this would basically tell you how many folded domains your proteins has right. The number of such jumps or whatever kinks that you have that would be equal to the number of domains in the folded structure of the protein.

So, these sort of single molecule experiments are very useful in order to see how a protein or whatever macro, whatever macro molecule that you are interested in how that has sort of folded in three dimensional space. Yes.

Student: (Refer Time: 06:12).

How will you try to go back if you leave the force, if you leave the force alone?

Student: (Refer Time: 06:23).

Not, I do not know if use its not a reversible process of course, but if you leave the force alone and if the conditions are right; the protein will again fold itself unless you have completely broken some bonds or something, the protein would fold back into its native state.

Student: (Refer Time: 06:42).

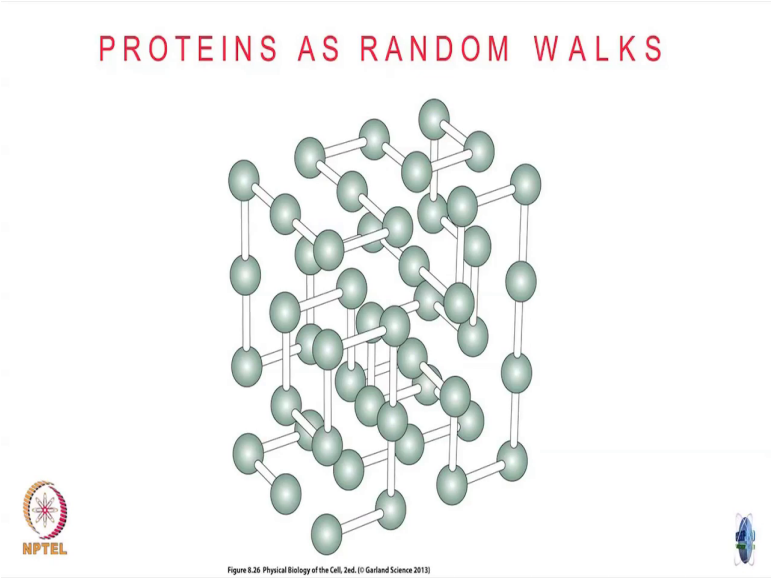
Yes, but it may not follow the same path right, so, in that sense it is not reversible, but yeah; it will it will fold back to its native structure provided the conditions are right for the protein (Refer Time: 06:57).

Student: (Refer Time: 06:59).

You can do it on the same protein yes and you should see a similar, you should see a same force extension (Refer Time: 07:04). Not really it is that the proteins, the domains in proteins are much stronger. So, if you just had one or two base pairs in RNA you would probably not even pick it up in a signature, you would need to have a relatively strong ish loop or not in order to see that. But, that is not something generic; you could have multiple here, there is nothing stopping RNA from having a; it is just that in proteins the domains themselves are much stronger because of the folded nature of the protein.

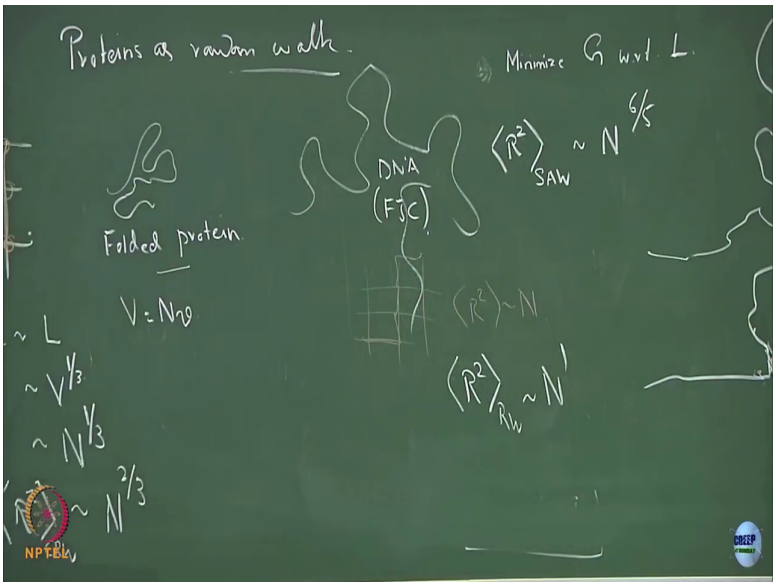
Since, we are talking about proteins, let us just, let us try to see how to sort of look at structures of proteins. We have looked at this random walk models, you have looked at these random walk models, that works sort of well for DNA and so on, but proteins are a slightly different beast. Precisely because they fold into these native structures, a random walk model does pretty badly with proteins or standard random walk model that is pretty badly with proteins.

(Refer Slide Time: 08:21)



What you could do for example, so this if you want to think of proteins in the context of a random walk, so proteins is random walk.

(Refer Slide Time: 08:36)

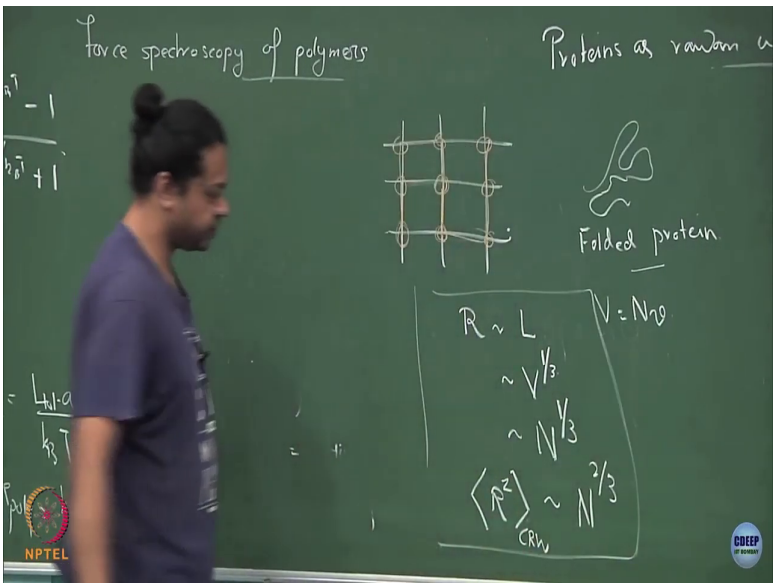


You have to take into account the fact these proteins are generally found in that folded native state, they have some sort of a conformation which is not random. So, they are compact objects, they are folded according to their sequence into the native state unlike DNA or some polymer like this it will just adopt some random conformation. So, DNA would be something like this, and this I could describe by a freely jointed chain. On the other hand for the folded state of a protein, you can use a different sort of a random walk models.

So, this is a folded protein, which is called a compact random walk. So, which is called a compact random walk, and the idea is that a protein is sort of going to fill up an available volume. So, not an available volume, but let us say, let me first define a compact random walk and then I will say. So, compact random walk is let us say I will define it on a lattice.

So, compact random walk is one, it is a self avoiding walk and it is a walk such that each of the lattice points of this space of lattice points is filled by one monomer or 1 unit of this random walk.

(Refer Slide Time: 10:08)



So, let us say if I had a lattice like this, if I had a lattice like this, this that is a very bad, something like this where I place a monomer on each lattice point is an example of a compact random walk. Note that this is so, this is space filling so each lattice site has one monomer associated with it.

And then you could look at different confirmations of such space filling walks. This is very different from the standard freely, the standard random walk that we are used to where if I

take a, if I take some lattice definitely not all lattice points will be filled; it will go as R square will go as typically as N.

On the other hand for this compact random walks, for these compact random walks because these are space filling the typical size will grow as the dimensions of your box the dimensions of your lattice itself which is nothing, but something like V to the power of one-third alright. And, this volume is not again because it is space filling this V is nothing, but the number of monomers times the volume of each monomer roughly.

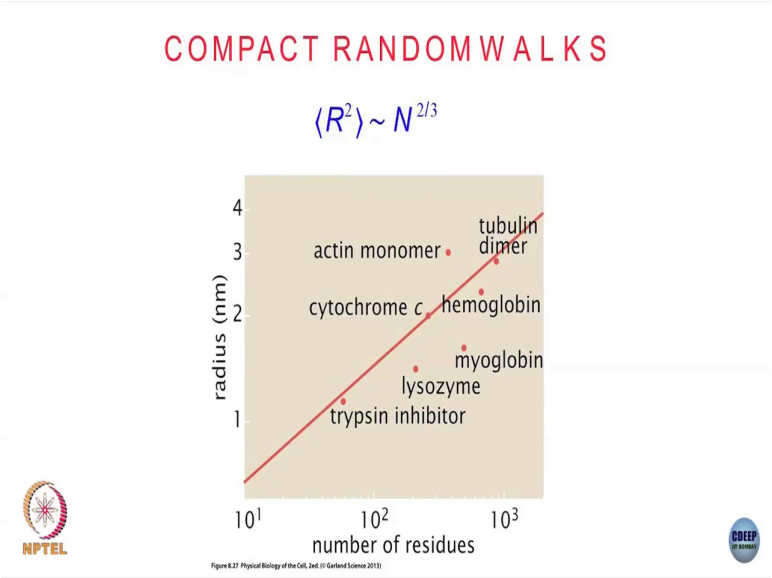
So, this is nothing, but R N to the power of one-third ok. So, this is of course, a very hand weaving argument, but you can show that this R square average grows as for this compact random walk grows as N to the power of two-thirds. As opposed to this; as opposed to this R square average for a standard random walk which grows as N to the power of 1.

Simply looking at this exponent tells you that this is a much more compact walk than this, this is much more open because the exponent is 1, this is much more compact the exponent is 0.67. And, just for reference R square of self avoiding random walk grows as N to the power of roughly 3 by 5 so 6 by 5.

So, this is even more open because this is self avoiding so, you cannot have intersections which sort of provides outward pressure if you will which forces the polymer to expand even anyway. For compact random walks this is the sort of scaling that one expects with this R square the typical size grows is N to the power of two-third instead of growing as N to the power of 1 ok.

And, this is the sort of paradigm that will use to model proteins, but let us just see how good this is. So, if you take different proteins of different sizes and so on and calculate the average size of these proteins this is what is.

(Refer Slide Time: 13:37)



So, this is the prediction of this compact random walk model, if you plot proteins with different numbers of residue spanning a couple of decades, well not a couple (Refer Time: 13:52). Somewhere over here is proteins like 60 70 amino acid proteins, here there are proteins which have 1000 amino acids and so on. And, you plot their observed radius as a function of the number of residues and the solid line is this N to the power of two-third line.

So, you will see that the, on an average this N to the power of two-thirds sort of a random walks so, this compact sort of a random walk does fairly well describing what these, what the typical size of the protein will look like. It is not as good as, what the and it is not as good as what this DNA, what we got for the DNA where it was very nicely N to the power of 1 and that is because sequence specificity plays an important role.

But, on an average this random, this compact random walk model is not too bad, you can say that given in the number of amino acids typically it will be somewhere around this N to the power of two-thirds. You may be wrong for a particular protein, but on an average if you did this prediction for many many proteins you would roughly be right ok.

So, we will what we will try to do is that we will try to use this sort of a model in a very very basic sense, in order to see what we can say about this protein folding problem finding the native state of a protein.