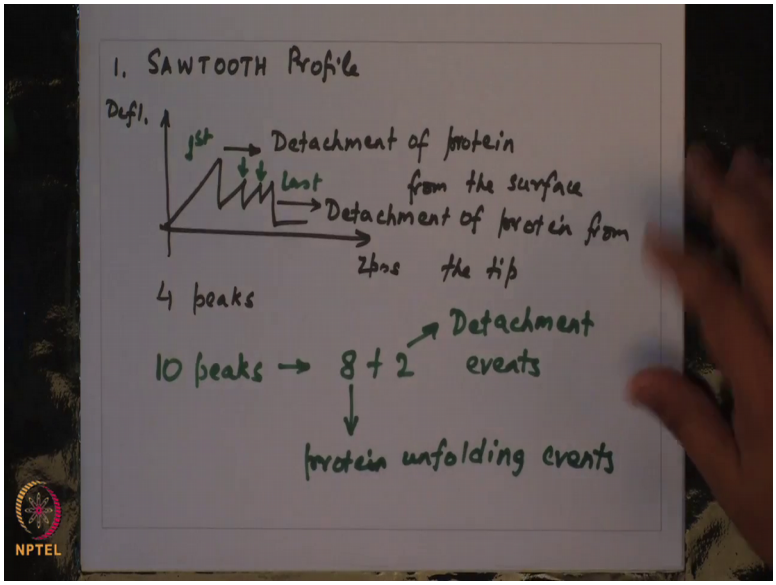
Introduction to Mechanobiology
Prof. Shamik Sen
Department of Bioscience and Bioengineering
Indian Institute of Technology, Bombay

Week – 02
Lecture – 10
Design of protein constructs for AFM

Hello and welcome to our 10th lecture of our NPTEL course Introduction to Mechanobiology. In last 2 lectures I have been discussing about, how we can probe protein unfolding using AFM and how we can use that for studying, how forces affect the unfolding dynamics of proteins and taking the specific example of phi printing.

Let we first start by you know doing a brief recap of what we observed, or what we discussed in last class.

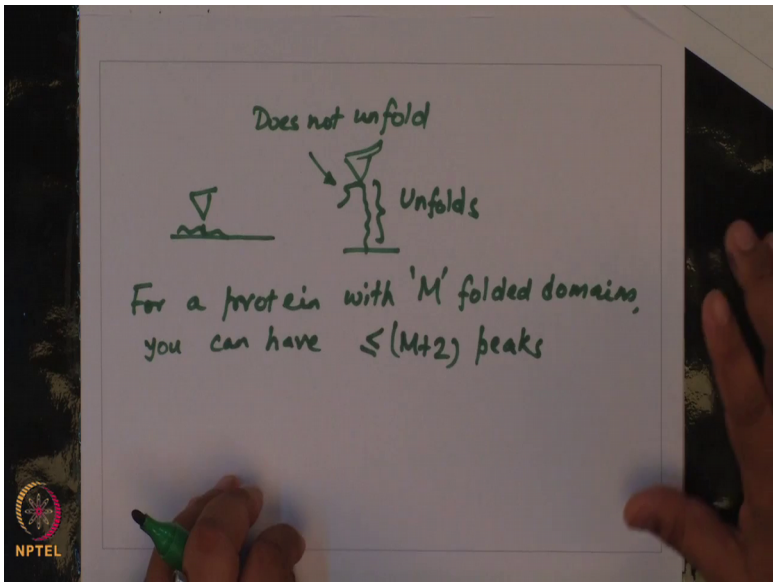
(Refer Slide Time: 00:50)



One I introduced Sawtooth Profile. When you do protein unfolding studies with sparse concentration and you pull a protein with an AFM. So your sawtooth profile looks something like this. In this case, what you see is there are 4 peaks. What happens how do we assign these peaks. So, I said that this first peak is associated with Detachment of protein from surface from the surface. And the last peak corresponds to detachment of protein from the tip.

Essentially the middle peaks these 2 peaks are your data peaks for your protein of interest. So, typically let us say you get a total of 10 peaks in a given experiment. Then you can bring it down you can break it down into 8 plus 2 because these two peaks corresponding to the first peak and the last peak you know are not protein unfolding events, but these are detachment events. But these 8 corresponding they correspond to protein unfolding events, but it is not necessary.

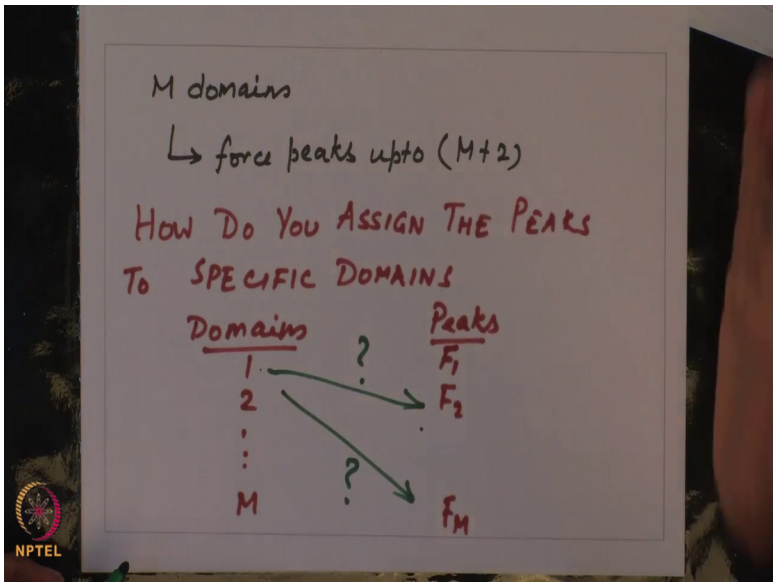
(Refer Slide Time: 03:15)



For example, let us say you did this pulling experiment and you got 8 peaks, but it is not necessary that you will always get 8 peaks. And the reason for that I explained in last class was, when your protein is absorbed you have no control, where your tip hits the protein. So, in a once the tip starts going up you a have an overhanging portion this does not unfold, only this segment, which is being stretched this is what this unfolds.

So, you can have for a protein with 'M' folded domains, you can have less or equal to M plus 2 peaks. So, at max you will have M plus 2 peaks, where the tip actually contacts the one end of the protein, but in all other cases you will have less than M plus 2 peaks. So, this is one thing to be noted.

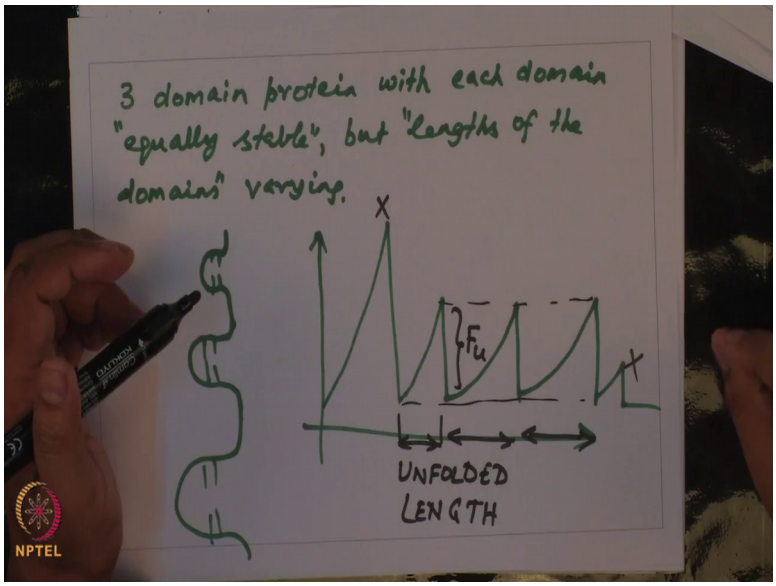
(Refer Slide Time: 04:37)



Let us say, find you got your peak right first thing is to record the data, but now it comes to analyzing the data. Let us say you have a protein with M domains, and you get force peaks which have up to M plus 2. So, you know that there are M peaks. Now, it comes the tricky part comes is, how do you assign the peaks to specific domains. In other words you had these domains 1 2 up to M and you have peaks F 1 F 2 dot dot F M.

You do not know whether 1 corresponds to let us say F 2 or 1 corresponds to F M. So, you do not know that. So, that is the challenge how can you assign the peaks to specific domains.

(Refer Slide Time: 06:04)

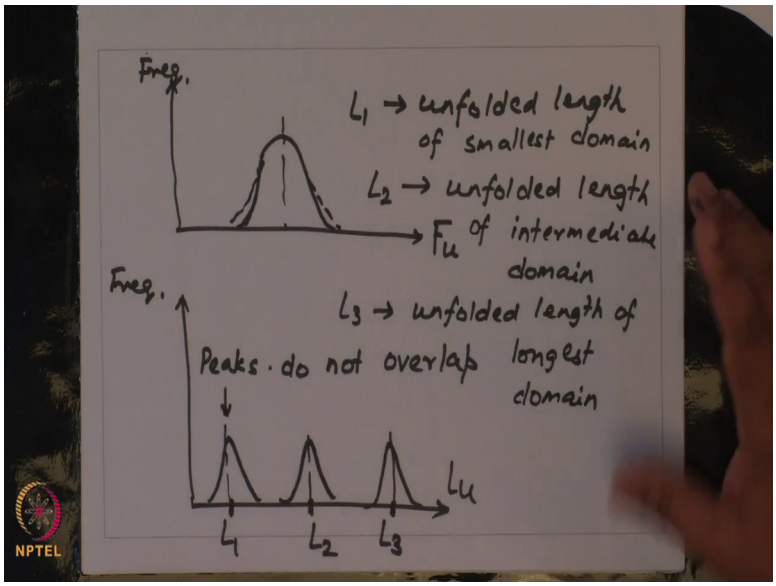


Let us take a simplest case you have a 3 domain protein, with each domain “equally stable”. So, equally stable would mean that the unfolding face is equal, but “lengths of the domains” varying. In other words, what I am saying is I am drawing these 3 domains.

The force required to unfold each of these domains remains the same, but unfold the sizes of these domains are varying lengths of domains varying means the domain sizes are varying. Now, let us see what the force curve the sawtooth pattern would look something like. So, once again I can dismiss this and these last 2 the first and the last peak I can dismiss, but I have 3 additional peaks.

So, 3 peaks, because what you see is the magnitude of these forces are the same, but these lengths are different. So, these lengths are different. So, what does this? So, for a single domain let say you have a peak like this. There are two things that you can determine from this peak 1 is force of unfolding and this is the unfolded length. So, your X axis gives us unfolded length and the Y axis gives you unfolding force. So, in this particular case because you know we took a case where each domain was equally stable. So, your unfolding forces are similar in magnitude, but the unfolding lengths are not.

(Refer Slide Time: 08:51)

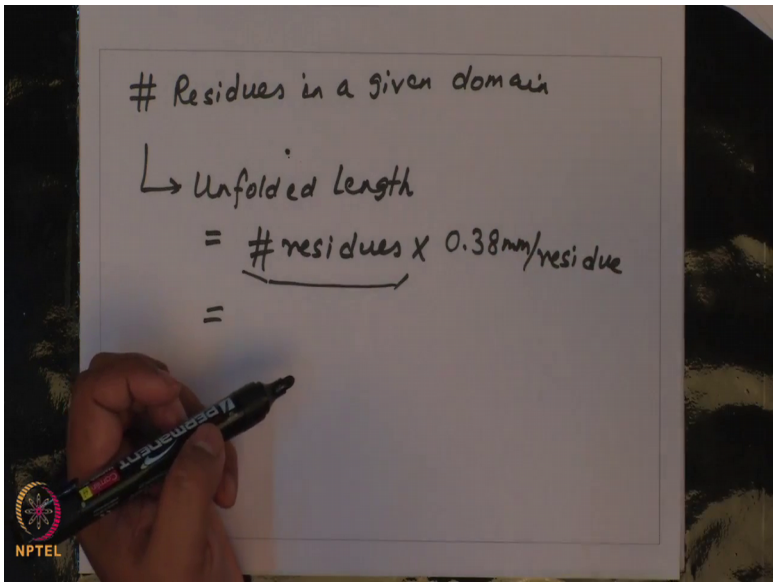


So, this is one such representative data point you have to do it multiple numbers of times and need to determine the statistics. So, what you plot F sorry, what you will have in this particular case that we drew. So, what I have drawn in the. So, I have drawn 2 histograms one is the unfolding force and one is the unfolding length. So, for the unfolding force because we assume that all the forces were of equal magnitude that the domains are equally stable I will get a single force peak.

It might have a little bit more scattered because you are collecting the statistics from, many events and there might be small variations in the standard deviation. But if these domains are of distinct lengths you will see that you will have 3 histograms, which do not overlap with each other or the overlap might be there, but the peaks do not overlap. You know, you have to take this value these mean values of each of the histograms and you have to assign this each of these lengths.

Let us say L 1 L 2 L 3. So, l one is the unfolded length of smallest domain, L 2 is similarly the unfolded length of intermediate domain and L 3 is unfolded length of longest largest domain.

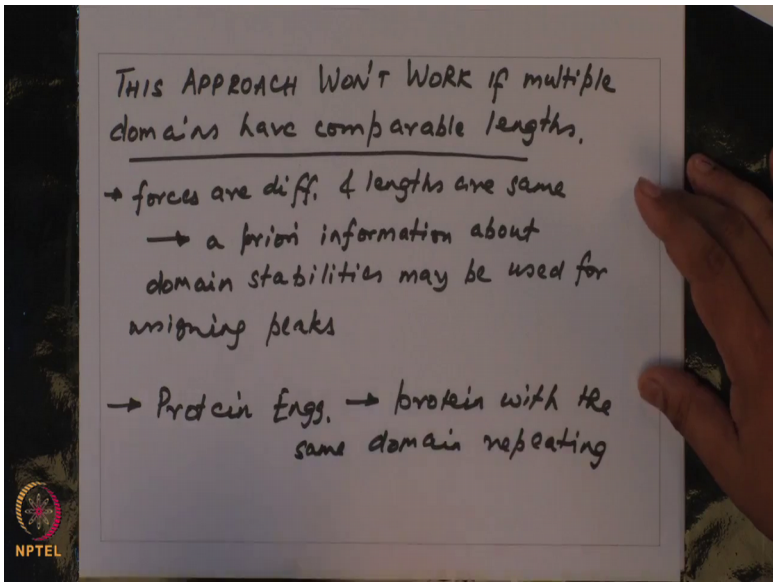
(Refer Slide Time: 11:47)



What is the unfolded length? If you know, if you know the number of residues in a given domain; then your unfolded length should be equal to the number of residues in 2. So, you should give you the number of residues into 0.38 nanometer per residue. This you can directly find out the unfolded length, now because you know the number of residues because you know the number of residues. You can calculate this. Whichever unfolded length that you experimentally measure tallies or is closest to this particular value, and then you know you can assign that this particular L corresponds to this domain. So on and so forth.

So, this is how you do the assignment for the simplest cases in let us say you have engineered polypeptide of the same peptide or things like that, where you have the same repeating unit.

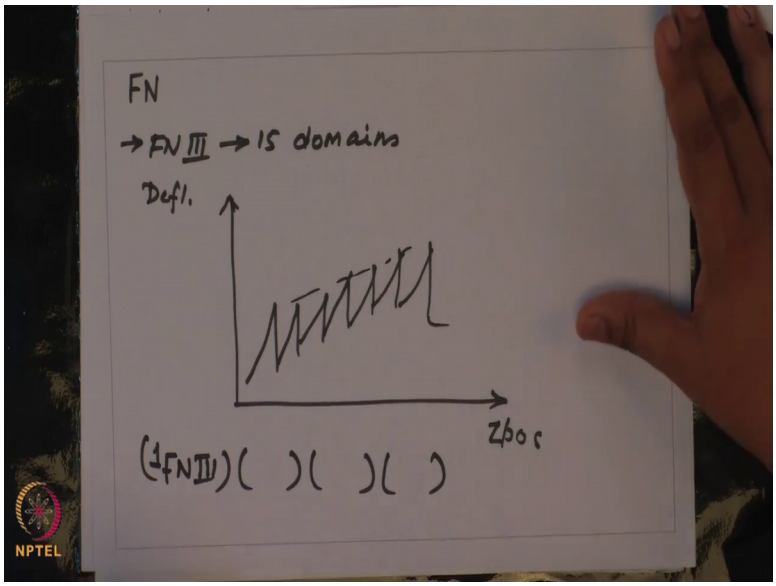
(Refer Slide Time: 13:15)



However this approach would not work if you have multiple domains. This approach would not work if multiple domains have comparable lengths. So, the simplest approach will even work for the other case where you are forces; if forces are different and lengths are same. So, if you knew a priori. So, a priori information a priori information about domain stabilities may be used for assignment assigning peaks.

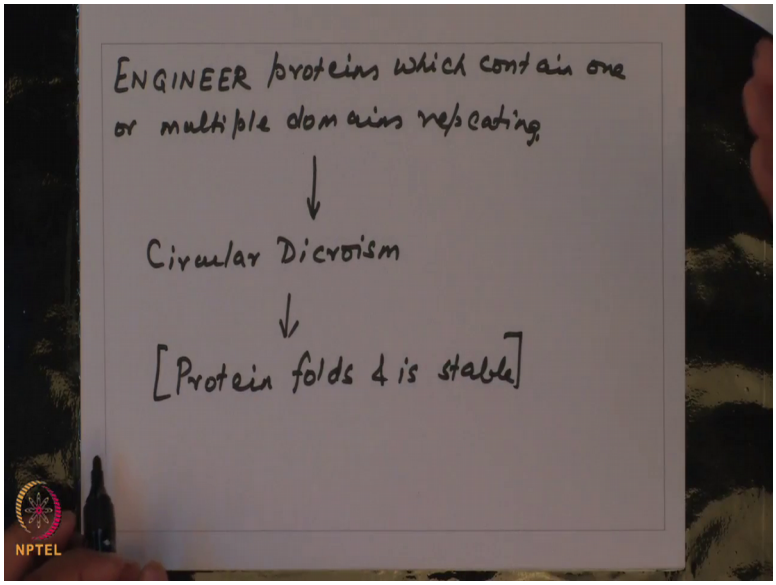
But clearly for this case this approach would not work if you have multiple domains which have comparable lengths. So, what is done in these cases is actually using protein engineering, where you make you know poly protein. You have a protein with the same unit same domain repeating.

(Refer Slide Time: 15:19)



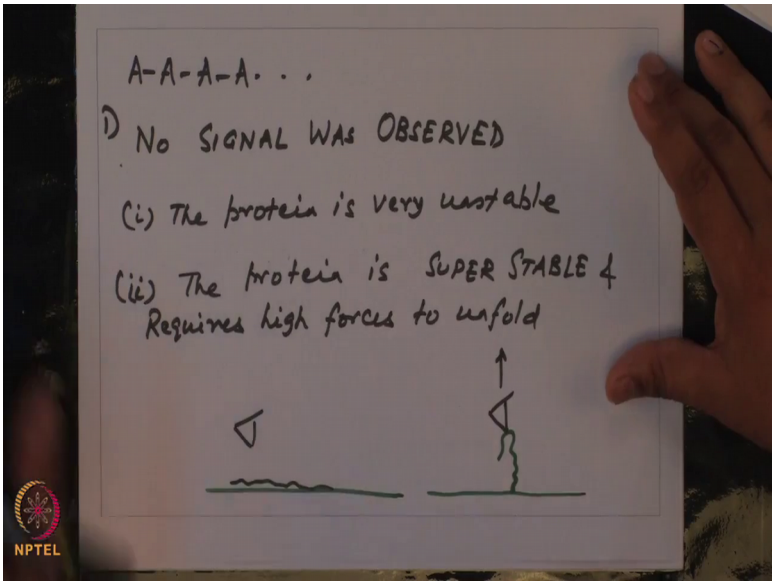
Let us take the case of fiber knitting. So, the FN 3 it itself has 15 domains. Now if you pull this long of protein and you have a c position if you pull this long of protein one thing that you observe, one thing that you observe is you will get these multiple peaks. So on and so 4th, I am not drawing all 15 domains, but you will have these multiple peaks, which might have comparable lengths, comparable process so and so, 4th. So, it is near impossible to assign the peaks. So, for this reason what you have to do is we have to construct protein, you have to engineer proteins, which have let us say domain 1 FN 3 this domain is repeated and so on and so 4th.

(Refer Slide Time: 16:43)



So, you prepare this you know you engineered this proteins. So, when you engineer these proteins. So, you have to engineer proteins which contain one or multiple domains repeating. Let us take one such case, now let us say you have made your protein. So, now one important step is after you have made this poly protein you have to make sure you use Circular Dicroism, and other techniques to make sure that the protein folds any stable is stable this you must check, before you do any AFM you must check that the protein that you engineered force stably, after you have done that check.

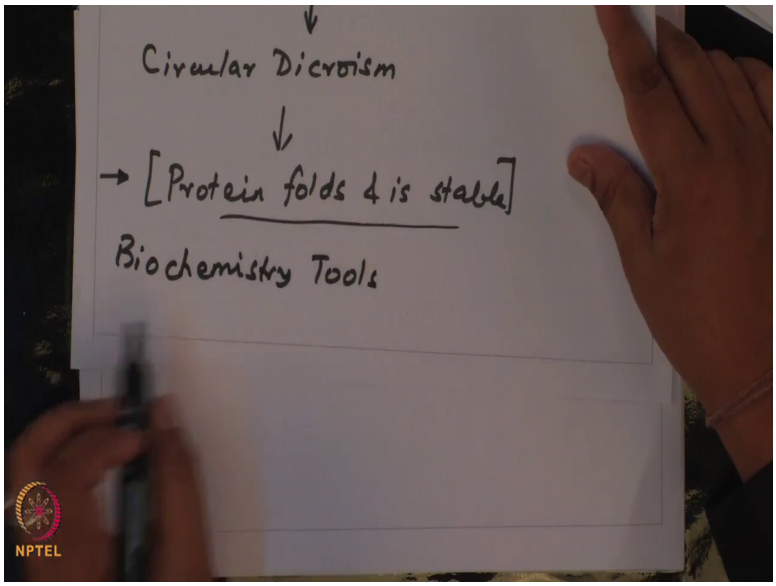
(Refer Slide Time: 17:54)



Let us see you have made a protein which has this unit repeating A-A-A-A so on and so fourth. So, typically if you choose a longer length, there is a greater likelihood that the protein will fold verses making just 2 ways. But let us say you did that AFM experiment you did the AFM experiment and you saw no signal at all. So, you have this situation no signal was observed. So, in this particular case what might be the possibilities? One possibility one possibility is let me draw. So, let us say that the protein is very unstable.

So, to begin with it does not fold properly and that is why I said that we must make sure that the protein folds and stable, but if you have done this check using biochemistry tools.

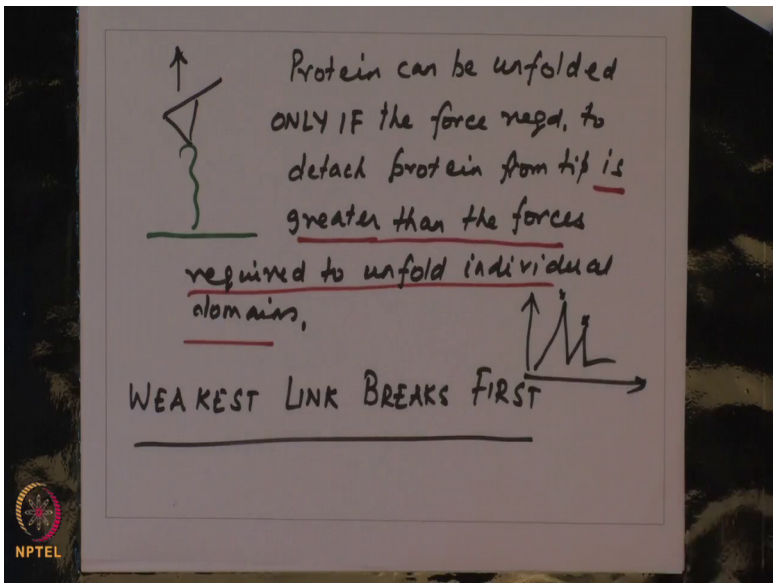
(Refer Slide Time: 19:00)



Then you know that this is not the case, but this is this might be one possibility that the protein is very unstable. What might be another possibility; the protein is super stable and requires high forces to unfold. This is somewhat counterintuitive why am I saying. So, if we know that the protein is stable why are we not able to detect it the reason is as follows.

So, you have your substrate and your protein is anchored right. So, when you pull up the protein. So, you reach this configuration you reach this configuration let us say.

(Refer Slide Time: 20:31)



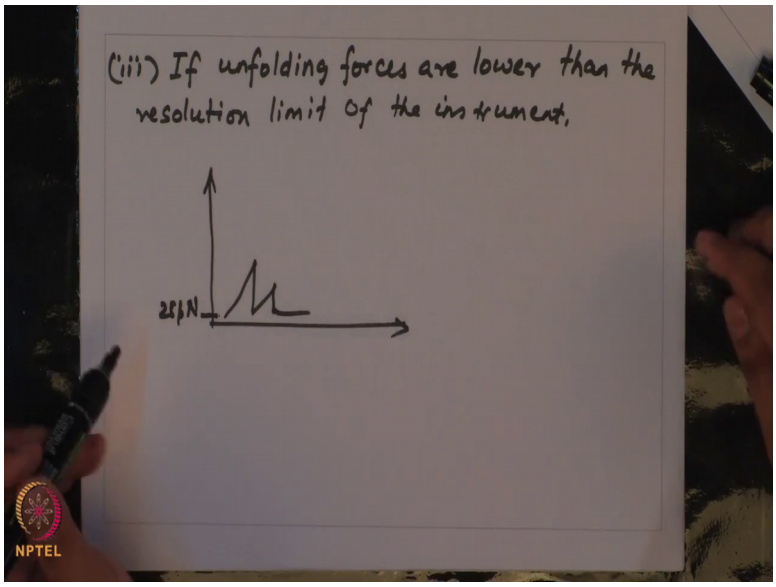
So, the tip is moving up and you are trying to pull the protein. So, they is the tip will be able to pull the protein only if that the force required to unfold upload a domain. So, this will work let me write it down here.

So, you have this situation. So, protein can be unfolded only IF the force required, to detach protein from tip is greater than the forces required to unfold individual domains. This is the criteria. Let us read it again protein can be unfolded only IF, when I say protein can be unfolded I mean, that the unfolding signature is observed with the experiment. So, protein can be unfolded only IF the force required to detach protein from the tip right. So, this is a system of springs connected in series.

So, it is always that the weakest link breaks fast, because the weakest link breaks fast. If the force required to unfold a domain is much greater than the force of addition between the tip and the protein, then the protein, then the tip reach is detached from the protein. So, you will not see any signal you will only see 2 detachment peaks. In this particular case here that is it this is all that you will observe. So, this is associated with detachment of protein from the surface and this is associated with detachment of tip from the footing.

In this case, so if you protein domains are very stable then also you would not see any signal.

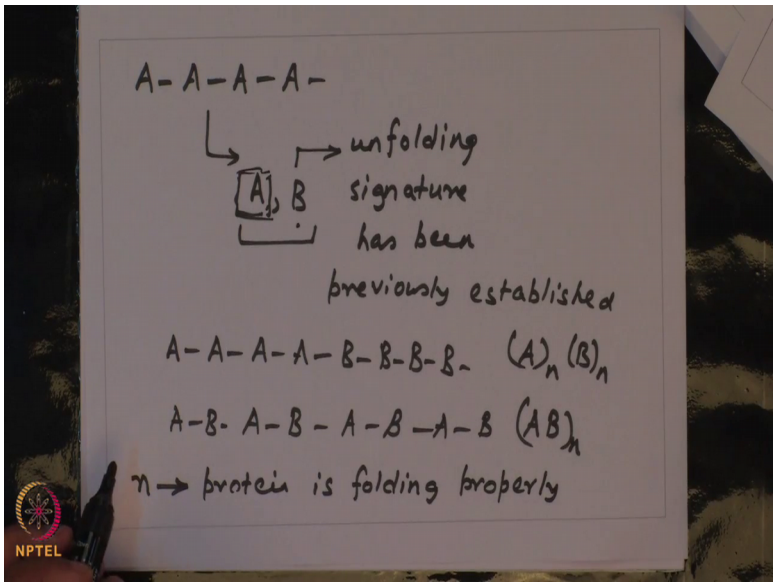
(Refer Slide Time: 23:20)



There are 2 possibilities we have seen; protein is very unstable then you would not see much, protein is super stable and requires high forces then you will see it. There is another situation. If unfolding forces are smaller are lower than the resolution limit of the instrument.

In other words in your 4th scale, if you cannot resolve below 20 pico Newtons. So, you do not see anything below 20 pico Newtons. This absolute lowest is let us say 20 5 pico Newtons. So, you cannot record you would not get any spectra for of peak, which is below 20 5 pico Newtons. In that case you will again see the same thing you would not even notice. So, this is the other case where the unfolding forces are very low then also your instrument you cannot detect those forces. But still there is a broad area over which you can still operate and find out what are the unfolding spectra.

(Refer Slide Time: 24:39)



Now comes the aspect of designing or protein construct. So, I said that you will make this poly protein. Now let us say you have not got 10 your signal you know that there are 3 possibilities. So, these 3 are the possibilities why you may not be getting the signal. So, there is very it is very difficult to find out which of these case applies to your protein.

So, what is typically done to get around it, but you know that your protein is folding stably as per other experiments. So, what you do here is instead of just having a poly protein like A-A-A-A you make a poly protein with A and B 2 components where B is a component, whose unfolding signature has been previously established. The rationale

being that if you are doing your experiment then you know that because B has an unfolding signature even if you do not get A then you know that your B signal should always appear in your protein unfolding signature.

So, in that case you can as in you can say that probably A has very low unfolding forces, that is why we cannot detect it or, that is one possibility. Now the next question comes is how do we make design our multi protein should we have a design like A-A-A-A is followed by B-B-B-B or A-B A-B A-B A-B. So, in shorts do I repeat A n times B and times or AB n times. So, n may be variable. So, you choose n such that at least you know that the protein is folding properly. So, your n might be 3 in one case or in might 4 in one case, but you need to make sure that this protein is holding property.

So, with that I stop here for today. So, we have seen what are the challenges associated with trying to trying to ascertain the folding signature of proteins and assigning it to appropriate domains next class we will continue from here.

Thank you.