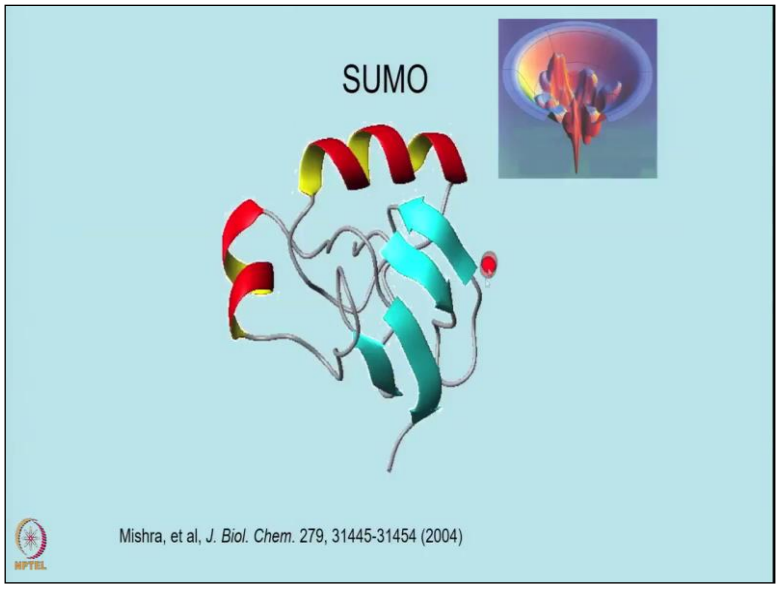
NMR spectroscopy for Structural Biology
Prof. Ashutosh Kumar and Prof. Ramkrishna Hosur
Department of Chemistry
Indian Institute of Technology - Bombay

Lecture: 38
NMR Analysis of Protein Dynamics - II

(Refer Slide Time: 00:29)



So we will continue the discussion with the protein folding and to introduce this protein folding phenomenon. So, I have here a particular example, we will deal with this particular protein called SUMO. This is a small ubiquitin modifier protein, is called SUMO, and what is shown here is the so-called folding funnel or it is energy landscape of the protein. What are the things, which are very important to understand this picture?

This bottom portion represents here, this is the native protein, which is very well folded, and on this axis, we have the energy of the system. And what is on the surface here? This is the model, which is called as the funnel model, what is here is the unfolded state of the protein. You see the funnel is very broad on the top, what does it mean? What does it indicate? It indicates different degrees of freedom.

This is the conformational heterogeneity is quite a lot. So, that is all represented by the width of this funnel. So, it is so wide here there are trillions of conformations for any particular protein. See how many conformations are possible for a particular amino acid. Let us say I

have a 100 amino acid residue protein and each amino has a two degrees of freedom the Phi and Psi.

Phi-Psi two torsion angles and each amino acid can take this any of these combinations of this. Therefore how many degrees of freedom will be there? 2 to the power 100. each amino acid is 2 degrees of freedom. So, the 2 to the power 100 conformations are possible for this particular protein, 100 amino acid residue protein, and the proteins in general fold relatively faster. If the system has to scan through all these conformational degrees of freedom how long it will take? It will take the age of the universe.

2 to the power 100 confirmations if it has to scan through and then finally reach a particular folded state that is an enormous amount of time. But nonetheless we do see that the proteins fold in much smaller time scales, sometimes minutes, sometimes seconds, sometimes hours, so on so forth. So, this was actually called as the 11th house paradox. How does the protein folds? How does it scan through the conformational space and reach to the particular ground state or the lowest energy state?

Does it pass through the entire energy landscape? So, this is the energy landscape of the protein. So, here on the top of the funnel, you have these trillions of conformations which are possible they are all interconverting very rapidly. And if it has to fold, you have to go down the funnel losing the degrees of freedom slowly. As it starts to fold some degrees of freedom are satisfied and some ratings of freedom remain.

And eventually it has to reach here. But the path it can take from this point to this point or this point to this point can be very different. So, there can be multiple paths for the folding. Because it has to eventually reach the same native state, because the native state is supposed to be very, very narrow, it does not have too many conformations, it is one confirmation.

So, we know that, that is what is called as the folded protein and this is the one which is responsible for the varieties of functions of the protein and we try to understand that. Now as the protein is going down, it can have different paths. So, if I want to plot here the energies of this. Suppose they have an unfolded state with a certain energy here which is high energy and if it has to go to the folded state it has to come down to the folded State here.

This is the folded State let us say. This is the folded state energy, this is unfolded, does it come directly here like this or does it come like this, like this, like this, and these are the various barriers and at every step, there can be a kinetic process is operating. Which is the one which is going faster? If the barrier is very high then actually the protein may get stuck in any of this minimum here.

There can be local minimum here, a local minimum here, then if it gets stuck here is not able to come out of this minimum, then of course it will reach the what is called as the misfolded state. So, the system will have to come out because of such kind of situation that is what is schematically indicated here. There are various barriers indicated here and therefore there are local minima there in this folding funnel.

So, depending upon that so the particular amino acid can fold go over this applies to every amino acid residue because the folding has to happen for all the amino acid residues right. Is there a cooperativity in this? If there is a amino acid residue which is folding does it induce the next amino acid to fold in a particular manner. So, is there a cooperativity in that? What is the energy landscape for every amino acid residue to go down the funnel and come down here?

So, these are the various intermediate states that can happen there. Eventually it will have to pass through different intermediate states and finally reach a particular folded state. If it is able to reach one particular folded state then we say it has the unique pathway and all the molecules are coming down to this. But if a particular segment is such that it is not able to come down to this state it gets trapped somewhere here then you generate a so-called misfolded state.

And these misfolded states can reel to all kinds of problems, they can aggregate, and they can lead to different diseases and things like that. Therefore, to understand the folding process is quite an important thing to do. Now so, how do we do this? We have to create certain situations like this. How do we understand what is the state of the protein at somewhere intermediate level here?

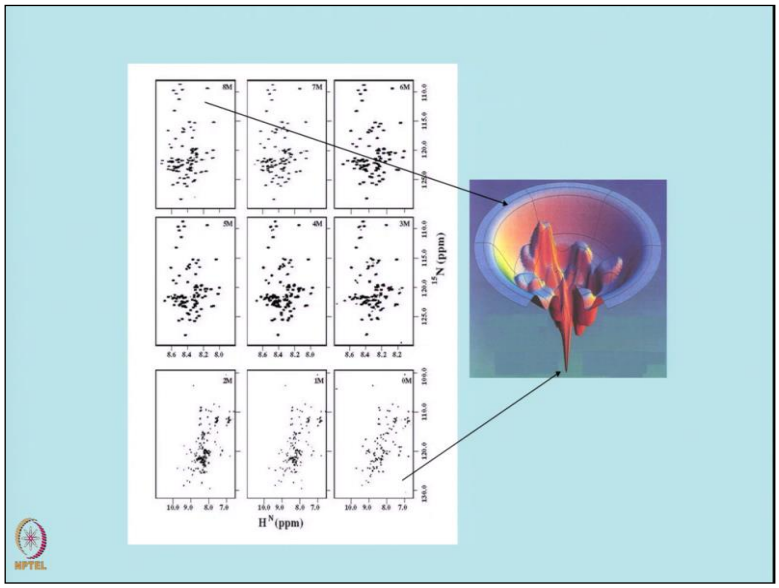
Ideally one can do it in a kinetic manner, but the kinetic manner to do that is you don't have the tools for doing so, appropriately at the residue level, but we can actually do the equilibrium states. At the equilibrium states, this also represent the conformation, this process samples equilibrium confirmations and it is going down. What are the equilibrium states that are present

at this if it passes through here I create a situation that the protein gets trapped somewhere here in this some of these ones.

The various equilibrium states I can create at different stages of the folding as it comes down from here to here, I can trap the thing at different stages, and these ones by doing artificially. So, this is what can be done by using the denaturants. By controlling the denaturant and concentration you can create intermediate equilibrium states, they may not necessarily be the kinetic intermediate states.

But they can be the equilibrium intermediate states, the protein will possibly go through some of those states as well when it has to reach finally these states the different segments. Because you notice here, the different parts of the unfolded states can fold to different pathways here to reach the final folded state. How do we do this?

(Refer Slide Time: 08:53)



This is what is indicated, this is a strategy. You actually denature the protein. First of all we have to create this unfolded state. So, you denature the protein artificially by adding 8 molar urea. Notice here when denatures a protein the chemical shift dispersion on the amide proton. These are HSQC spectra amide proton here and this is the N15 chemical shift here and you are plotting here the HSQC Spectra at a different urea concentrations.

This is 8 molar area, 7 molar urea, 6 molar urea, 5, 4, 3, 2, 1, and 0. This is the final folded state, that corresponds to this state here and this corresponds to this state here, and notice here the chemical shift dispersion is very very small and that is indicated. So, 8.0 to 8.6 only 0.5 to

0.6 ppm is a dispersion here. That is why these peaks are looking quite broad in this area. You compare that with the situation here.

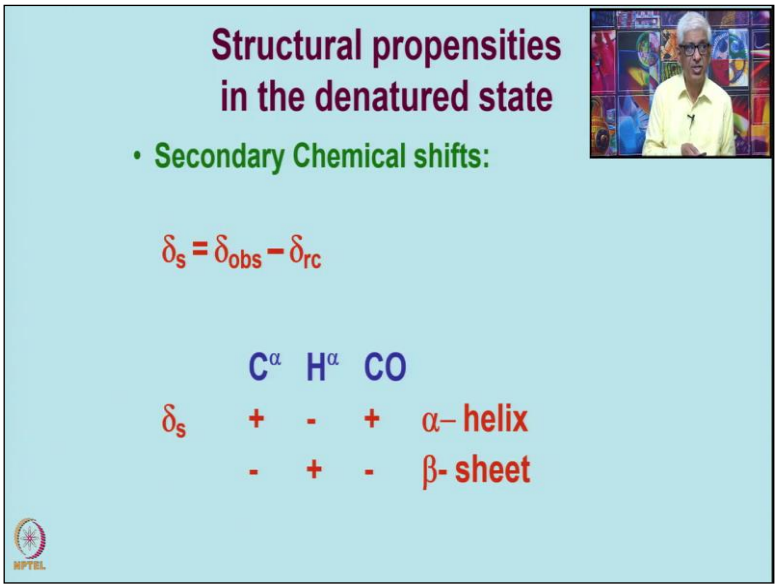
This is 7 ppm to 10 ppm, this is 3 ppm, whereas this is only 0.5 to 0.6 ppm. So, therefore the peaks are looking quite bigger here and that is just because of the scales. 8 molar urea you have this very narrow chemical shift dispersion here. The N15 dispersion is of course good. But now you see we had these tools which discussed last time with regard to handling unfolded proteins, the IDPs, intrinsically disorder proteins or unfolded proteins.

So, there you can use this HNN-HNCN spectra to assign the individual peaks. Now you have to assign the individual peaks in every spectra, because there can be differences here in the chemical shifts depending upon what the 8 molar urea is doing. Because the urea goes and interacts with the backbone of the protein, therefore there can be differences in the chemical shifts. You will have to identify individual residues in each condition.

What is changing from here to here? The chemical shifts will change, intensities will change, the relaxation properties will change, as you go from down the funnel like this. As you go down, as I said, the number of degrees of freedom will start decreasing. As you go down here, you have the maximum degrees of freedom. As you start folding, some structures that are getting formed which will restrict the degrees of freedom.

Therefore, the number of degrees of freedom will start going down as you go down the funnel here. So, therefore as you move down the funnel here, so, you have the peaks, some structures are getting formed, some areas are getting mobile, and that can be seen by the relaxation properties of the individual nuclei in amino acid residues.

(Refer Slide Time: 11:53)



So, let us see, what is it we get in this? How do we measure the structural changes that are happening as we are going down the funnel of the protein? There is one particular parameter, which is extremely useful that is known as the secondary chemical shifts. You know that in the particular protein the chemical shift indicates the environment around the particular nucleus. If the protein is folded it has a particular kind of an environment, if it is unfolded completely random coil it has a different environment.

And the random coil is more like an individual amino acid residue, individual amino acid residue does not have any structure. Therefore, the random coil conformation can be taken as the individual amino acid residue. So, if there is a chemical shift which is different from the random coil chemical shift, it indicates certain degree of structure around this particular amino acid residue. For any particular nucleus if you are monitoring, it shows a certain amount of restriction or the conformation preference around a particular residue.

So, you define what is called the secondary chemical shift δ_s,
which is $\delta_s = \delta_{obs} - \delta_{rc}$
This random coil is basically an individual amino acid residue. Here of course, the neighbouring residues also will have an influence. So, typically one generates a table for at every particular amino acid residue keeping different neighbours and you get a certain kind of a near neighbour effect etcetera.

All of that put together you generate a random coil chemical shift for a particular amino acid refer. So, we add corrections for all those neighbouring effects. Now and what are the nuclei

one can monitor? You can monitor whatever are the nuclei possible, you can monitor the N15s, you can monitor the amide protons, you can monitor the Cα, Hα, CO; and I have shown here a few of those.

Typically, these are the ones which are more sensitive to the secondary structures of the amino acid residues. So, here you have the Cα, Hα, and CO. These chemical shifts one has to obtain. I showed you earlier the N15 proton HSQC Spectra, but that alone is not enough. That will only give you the N15 proton chemical shifts. Having obtained those N15 proton amide proton chemical shifts you will have to also carry out other experiments, which we discussed earlier.

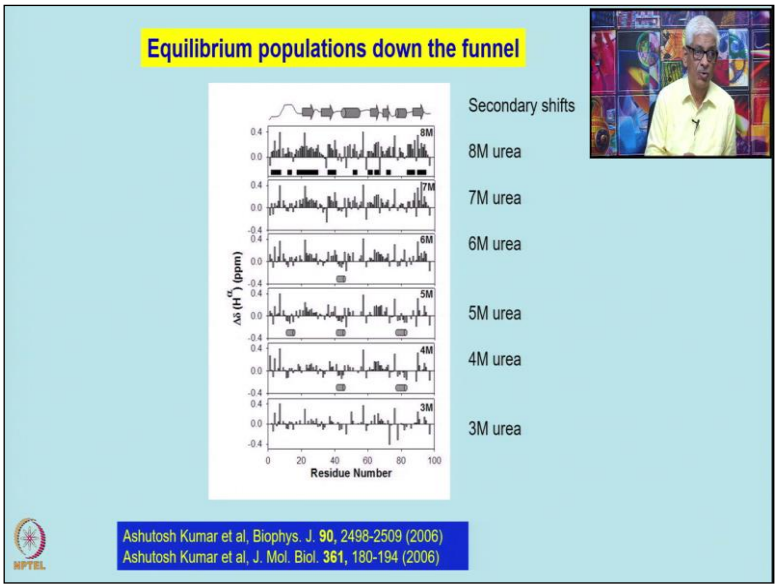
Like the HNCACB, HNCO, HNCA, all of these experiments one has to carry out to obtain the Cα, Hα, and CO chemical shifts. Now is there a trend in this for particular secondary structures? Yes there is a trend here. Now if you plot here, if you look at the secondary shifts for an alpha helical structure if there is a propensity. So, in a denatured State there is no regular structure but there can be a propensity.

What does the propensity meaning, it has a higher probability of obtaining a particular kind of a structure. It will not be permanently there it forms and goes, forms and goes, forms and goes. So, that is what is called as the propensity, is certain propensity for a helical structure. So, if we have that sort of a situation there, what will be the secondary shifts? The observed chemical shift minus the random coil shifts.

For the Cα, this secondary shift will be positive, the carbonyl also it will be positive, but the Hα will be negative. So, if a particular amino acid residue has a propensity to be in an α-helical conformation, then you have this combination of secondary shifts, whereas if the particular amino acid residues has the propensity to be in the β-sheet structure or the extended β structure is completely the opposite.

You have the Cα will be negative, Hα will be positive, and carbonyl will be negative. So, this kind of chemical shift changes you will observe in your experimental spectrum for all of these.

(Refer Slide Time: 15:47)



We will illustrate this for the SUMO protein and what are the secondary shifts, which are plotted here. It's in illustration. We have done that for various with along with the other nuclei as well Cα and CO, what is shown here is for the Hα. These are the secondary shifts. So, now notice here at 8 molar urea, this is the structure of the folded protein shown on the top here.

This has a long tail, N-terminal which is unstructured, you have a β-sheet, a β-sheet, the Helix, small β-sheets, β-sheet, Helix and the β-sheet again. This is the sort of a structure which is present in this. Now in 8 molar urea, what sort of a preference do you have? These are all positive. So, most of these are positive here, most of this Hα secondary shift, these are all positive and what does that indicate?

That indicates that you have a β-sheet propensity. So, most residues prefer to be in an extended conformation like in a β-sheet at 8 molar area. These ones indicate this small rectangles here these are the areas you can see continuously a range of 4 5 residues if you find then you draw a kind of a bar here.

So, those all β-sheet structure. Now this is the beta sheet structure, this is not present in the native protein here, this is extended. And here you have a β-sheet structure once more, then you have the β-sheet structure, this is present in the native protein β-sheet structure, this is present in the native protein partly here. And all over the place where there is a helix here, you do not find any helical propensity there.

You have a β-sheet propensity again, very small region and other regions here the β-sheet everywhere β-sheet. Where there is a helix you are not seeing the helical propensity there. Now you decrease the urea concentration go down to 7 molar. Now you see the changes that are happening here, see this all these with that big positive.

Slowly, there is a decrease in the positive values, and this remains everywhere. There is a small decrease in the positive values of these secondary shifts. So, now you go to 6 molar. While the decrease in the positive value continues to happen, you start seeing some little bit of a continuous stretch of Helix here. A stretch of a helix therefore indicates that as a helix here.

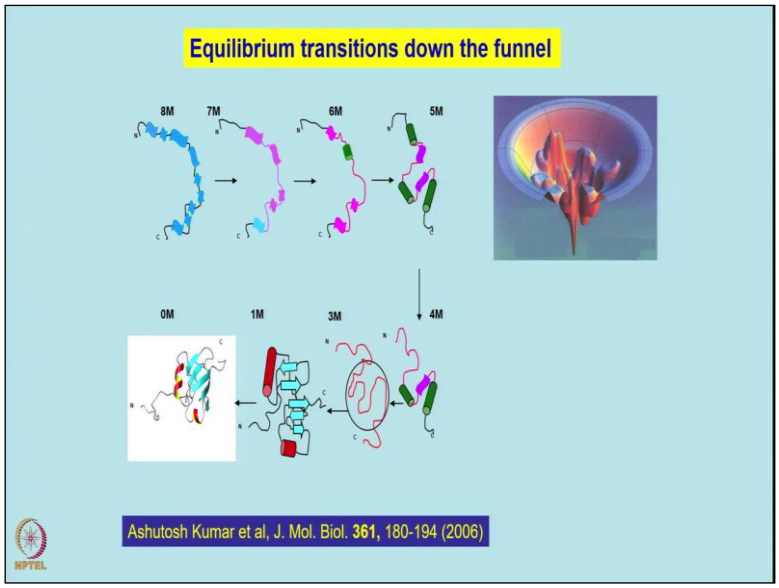
Little bit of the Helix is getting formed here, but rest of it is still β Although the propensity has decreased, there is a change happening in the structure but the propensity is still there. Because we said this is the probability at every amino acid residue, what is plotted is the probability. So, now you see there is a small Helix getting formed here.

Go to 5 molar urea, now you see small Helix getting formed here too, a small Helix getting formed here too, this continues and these numbers have decreased again, compared to this, these numbers have decreased again. Now we come from this to 4 molar urea, this is the same, this has decreased, this is the same, and this has vanished. This was about 4, 5 residues long here and that is vanished this is gone it is not.

So, much there and you come down to 3 molar urea most of these helices are gone, is all negative or small things they know structural preference it appears there is no structural preference at all here the propensities are not there. What does it indicate? Initially when we went from 8 molar down to 3 molar, let us go slowly, the initial extended structure propensity was there, this started decreasing and some helices were getting formed.

So, one Helix here and then 3 helical propensities not just this helices are not permanently stable, they are there, they are propensity. That is it is formed and getting broken, formed and broken, formed and broken, that is the way it is happening. And therefore, we reach a stage when you come to 3 molar urea, so, all these structural propensities which are there, these ones have gone. What does that indicate? This is shown in the next slide here.

(Refer Slide Time: 20:33)



So, here you can see this is the very dynamic 8 molar urea state and we come down to 7 molar urea, what is the colour change here? This colour change is the change in the dynamicity in the protein. There is a change in the relaxation properties. Why does it happen? Because some constraints, some structures are getting formed, because of that the dynamics in the protein is getting reduced, is not as flexible as it was in this case here.

So, as you go down to the 6 molar, you saw a formation of a helix here and the dynamics is reduced further and you see is going all the way until here also the dynamics is getting reduced, because some as you form some structures it introduces certain constraints for the protein to follow. Therefore, that will reduce the dynamics in the protein. This will get reflected in the relaxation properties especially the R_2 values, which we have not shown here.

But this information comes from the relaxation measurements. So, you have this reduction in the dynamics as you go down in this here. So, from high frequency dynamics, you come down to low frequency dynamics, because motions are getting constrained. So, that is millisecond you come down to the micro second to millisecond timescale, you know the R_2 values go up.

Now you go to 5 molar here, as I shown you, you got 3 helices formed and the constraints are still there, the motions are reduced. As more and more structures are getting formed, protein is getting more and more restricted with regard to the motions and that will be reflected in this other portions of the protein as well, because the protein is continuously changing.

You go to 4 molar, see this Helix has disappeared, as I indicated there in the previous one, this Helix has disappeared and you have these two helices are still there and this one is still there but overall the motions are constrained, that is a dynamic is getting changed. And now you come to the 3 molar everything is vanished, why is it so and this is very important, why is it so?

Because now if I look at this residues, where there is, on the basis of the relaxation measurements, it turns out that the protein is undergoing a certain kind of a dynamics, which will eventually lead it to the proper folded state. What is the folded state, that is here. This is the folded state, this is the original zero molar and the one molar structure is here. Because it has to form so many β-sheets, and then it has to form this helices at these two points.

But in none of these states these helices are there and none of these states you have these β-sheets permanently formed. Therefore the protein has to change from these states into this state. So, therefore what we call these as the non-native contacts. Initially the protein, when we went through a kinetic process of preferential structure adaptation, and then it forms certain structures, which were kinetically favourable.

Therefore, it formed some non-covalent non-native contacts. And these non-native contacts will have to be removed. So, if for that purpose the protein got prepared into a this state where this all these non-native contacts have been removed. Now it is all set to fold into a state where it is more like a native protein. So, when you go from 3 molar to 1 molar,

you see start seeing some β-sheet formed here, some helices formed here, and these are the native contacts and that you can see here as well in this spectra. So, in this spectra, you see from 3 molar, when you come to the two molar, the spectral dispersion is increased. The spectral dispersion has already gone up, compared to this spectral dispersion, this was only 0.4 ppm, the spectral dispersion is increased quite a lot here. You see these ones are small tiny dots

and this is even more increased, better, the peaks are becoming even more sharper and you are starting a greater dispersion along the amide proton chemical shifts. And then you come to zero molar, therefore is all clean beautiful spectrum, every peak is separated and the spectral range has gone from 7 ppm to 10 ppm and that is what you can see as the structure is getting formed.

So, here you see that sort of a thing happening at this point. So, the structure is getting formed at one molar, β-sheets are formed and the β-sheets typically have a wider spectral dispersion compared to the α-helices. α-helices are typically within the small range of 8.0 to 8.5 ppm whereas the β-sheets goes all the way to 9.5 to 10 ppm and that we saw in the in the HSQC spectra as well.

And this is reflected in these changes in the dynamics in the protein and the secondary shifts and the helical propensities form in the protein as we are going down the funnel from here to here, and remember this are all average effects, these are not one particular molecule. We are talking about the propensities. When you say propensities, it is the average effect.

So, therefore on the average, you have no structural propensities at this point and then you go to the one molar you have the propensity increase. Now they are stable structures getting formed, when the stable structures are there you have clean chemical shift dispersion and you have a wide range of chemical shifts reflecting the β-structure formation and then of course when you go to 0 molar,

you have the proper helices formed, the β-sheets formed and this is how the protein is getting folded. Therefore what is the message from here, the message is the proteins when they fold they follow a kinetic pathway and they do not necessarily go through the native contacts. It is not that these structures are getting formed straight away as the protein starts to fold protein.

These helices are not getting formed straight away as the protein starts to fold, you provide these conditions for folding it does not necessarily go to these states. It goes through whatever is feasible in the kinetic pathway. So, these ones are feasible in the kinetic pathway and therefore they form helices is here. But these are not the helices that have to be there, but as you provide further and further folding conditions these conditions these are not the most stable states.

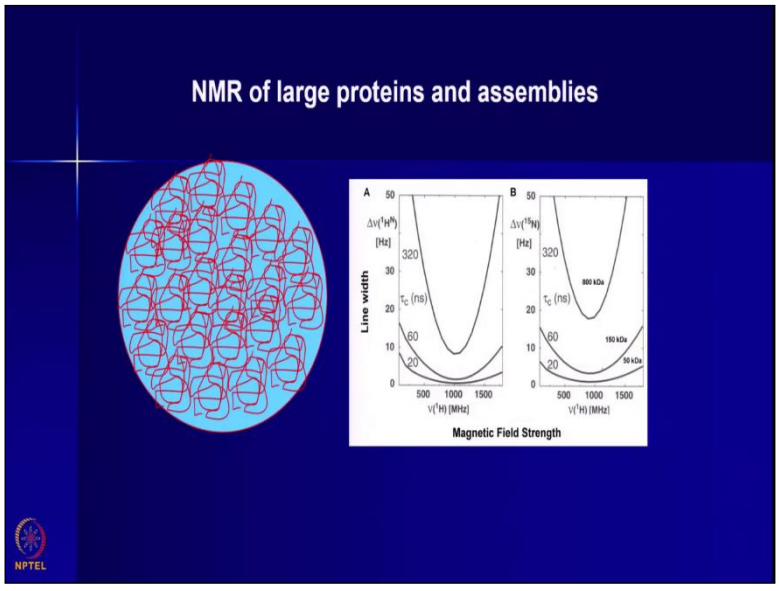
So, therefore these are temporarily formed and as they are providing better folding conditions, the eventually these native contacts will have to be removed and these ones are getting removed, system is prepared. And you see there is exchange rates appeared in this situation, that there are residues which come for come close and go, close and go, close and go, like this when it is happening the regions, which are coming close by.

They are making contacts, you will see larger changes in the relaxation rates there. The intermediate exchange will increase it is reflected in the R_2 values and that indicating that those contacts are similar to what are present in this state. Therefore, the relaxation data, there was indicating that these structures are getting formed. Because we saw changes in the relaxation properties of the molecule indicating that such a structure is getting formed.

So, once when this is formed, then of course you go to zero molar, you have this sort of a structure which is formed. So, this is how you can study the protein folding phenomenon by following the equilibrium transitions there. So, these represent the changes that are happening down the funnel and it is an average effect, every residue can follow different pathways, but of course in every molecule, you have this changes happening in different ways.

But what you see is the ensemble average. On the ensemble average, what you are seeing is what is reflected in this way. So, this is the great application of NMR for studying the protein folding phenomena in any particular protein. This is a small example but of course, one can do this for large proteins as well. Now so, I think this is we will go into the next class with NMR of large proteins and assemblies.

(Refer Slide Time: 29:41)



And this possibility of studying the unfolded states allows us to study aggregation phenomena in proteins. I think we will stop here we will take up this in the next class.