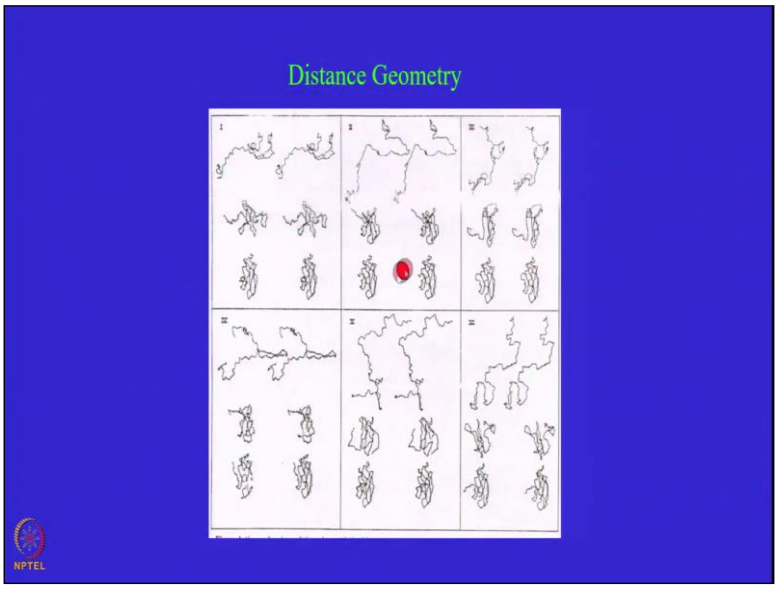
NMR spectroscopy for Structural Biology
Prof. Ashutosh Kumar and Prof. Ramkrishna Hosur
Department of Chemistry
Indian Institute of Technology - Bombay

Lecture: 35
Determination of Structure and Dynamics of Proteins - 5

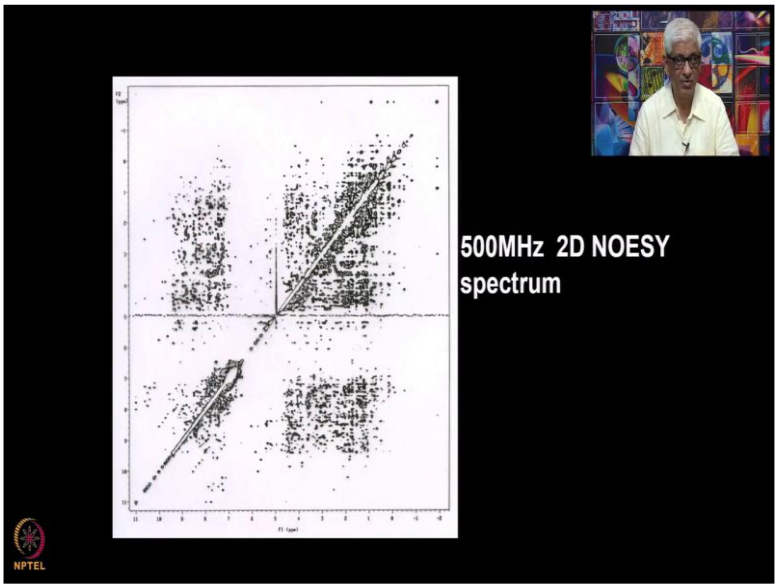
So, we stopped at this slide last time: how to calculate the structure of the protein using this algorithm of distance geometry. You have a set of distance constraints derived from the NOESY spectra which will tell you what a particular distance should be, within what range it should be. So, you have collected let us say about 1000 such kinds of distance restraints, then you build a model in your computer and is an algorithm called distant geometry which will optimize the various torsion angles along the polypeptide chain and you can get the final structure satisfying the distance constraints you might have.

(Refer Slide Time: 00:58)



So, that is the kind of an indication which I am just repeating. What I said last time, is that you start with the different initial structures for a particular protein and you have a set of distance constraints and as the computer starts optimizing it, eventually you will reach this final structure here. And the final structure is the same in every case regardless of where you started from therefore that adds confidence to the fact that your distance constraint set is complete and it is adequate to define a unique structure for a particular protein.

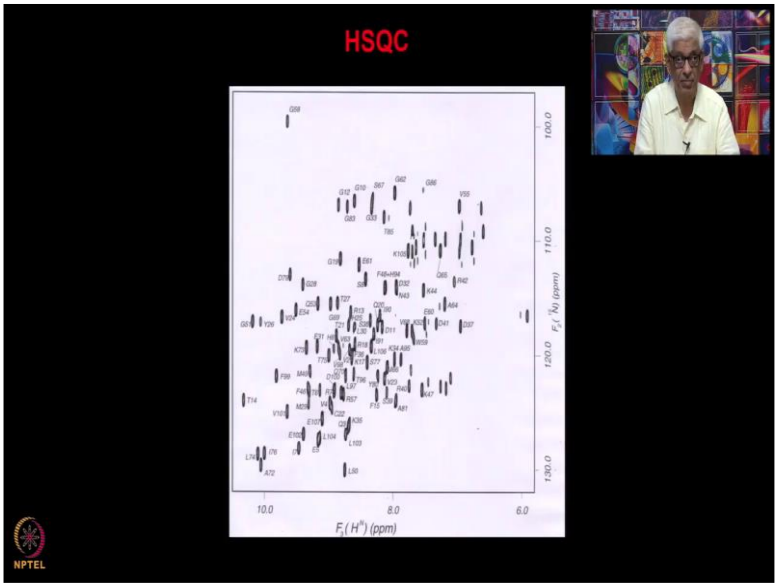
(Refer Slide Time: 01:30)



Well that work very well for relatively small size proteins like about 60, 70, 80, and 100 amino acid residues, but the moment you go to even larger proteins, then you start getting into difficulties. Because the NOESY spectrum is extremely crowded. See this is a spectrum, which has about 130 amino acid residues and then you see the number of peaks which are present here. This is enormously complex overlapping peaks.

And therefore quantifying each one of these peaks here is going to be an extremely difficult task and there are many peaks which are close by here and then you cannot quantify these peak intensities to establish the distance constraints. Therefore you already started feeling the inadequacy of the 2-dimensional NOESY spectrum. Now what do we do? This is all proton-proton spectra.

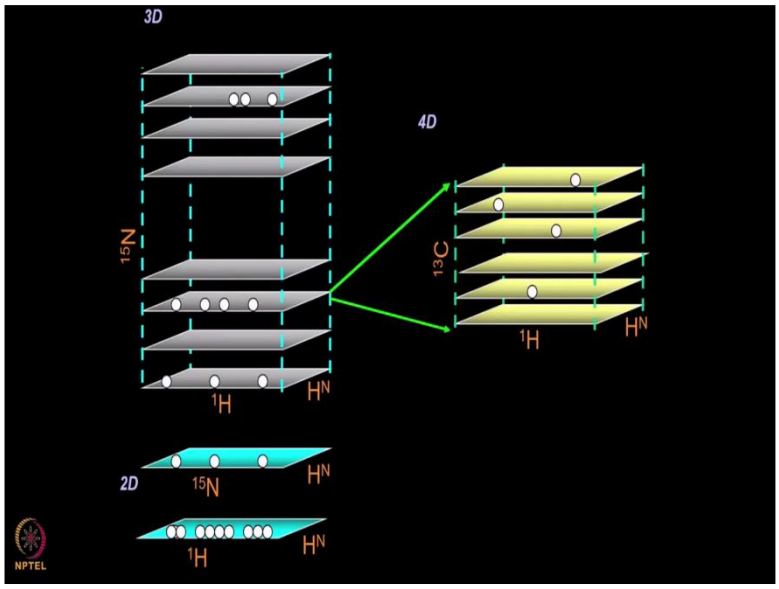
(Refer Slide Time: 02:22)



So, therefore what was done was to make use of other nuclei which may be present in your protein, in the sense that you have carbon-13 and nitrogen-15. So, here you have the nitrogen-15-proton correlation spectrum. So, you have these amide proton regions here, and these are the N15 chemical shifts of the backbone amides and here you can count the number of peaks.

Basically, we see correlation peaks here only, you do not see anything else. So, you have one peak per residue, each amide group produces one peak. So, there is a good dispersion of the N15 chemical shifts here and there is also good dispersion of the amide proton chemical shifts here. So, therefore you can make use of the N15 chemical shift as well to display your spectrum in more than 2 dimensions, then you can get more information from the NOESY spectra.

(Refer Slide Time: 03:10)



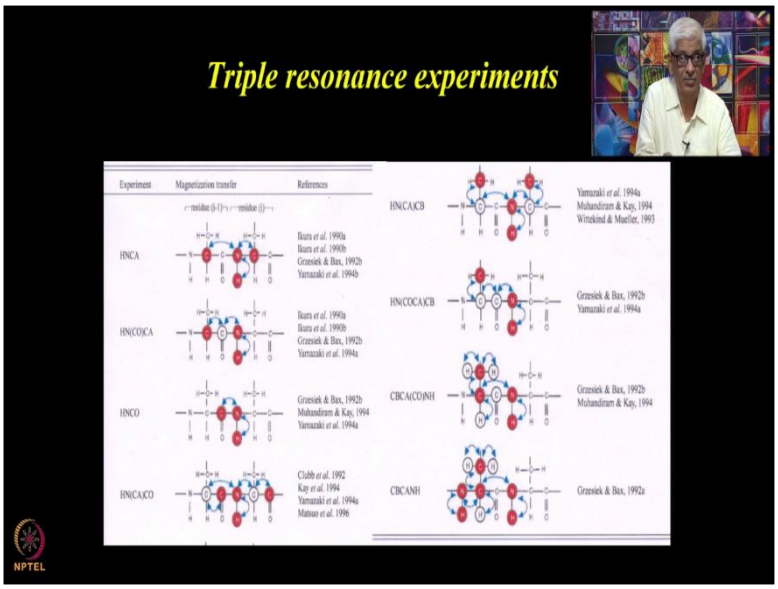
So, that is the schematic indicated here. So, you have the proton-proton correlation spectrum here, as we saw in the earlier spectra, then you also have the N15-proton correlation spectrum here, and then you would pull these peaks apart along the N15-dimension making use of the N15 chemical shifts here. There are three N15 residues, having the same amide proton chemical shift at this point.

Therefore when you pull them apart, then you will see they will get each one in a different plane, the three N15s come out in different planes. So, therefore the resolution is improved and you will be able to assign the individual peaks to individual N15 chemical shifts. Now to further improve the resolutions, you can go for C13 separation and make use of the C13 chemical shifts of the individual amide proton and then you can spread them apart on every plane, you have only one peak.

So, then this becomes easier to identify a larger number of peaks and quantify the intensities. So, therefore this may be a proton-proton NOESY spectrum and you spread them apart in the different planes, then you can quantify each one of those peaks, and then you can calculate the structure of a larger protein. So, therefore this means that you have to have N15 labelled in your protein, and you have to have also C13 labelled in your protein and this has to be done by recombinant molecular biological methods.

So, you express the protein inside bacterial cells *E.coli* cells and which are grown in the presence of N15-labelled ammonium chloride and C13-labelled glucose. So, the *E.coli* cells incorporate these nuclei in every protein that they synthesize and then after that, you isolate and purify your protein. Then you can study the individual proteins.

(Refer Slide Time: 05:07)



So, that is, therefore, you need 3 different kinds of nuclei and these are then called as triple resonance experiments, because they make use of proton, nitrogen and carbon-13. So, there are several sets of experiments which we have seen earlier when we were talking about the methods. So, we will now go back in a little bit of those details and try and explain how these things work some of those things will illustrate and the other ones can easily be followed thereafter.

So, you have this HNCA experiment. This is the triple resonance 3-dimensional experiment. So, you have the amide proton here and the nitrogen here and the 2 Cαs here. The magnetization transfer happens from the amide proton to this nitrogen, and from this nitrogen, it goes to this

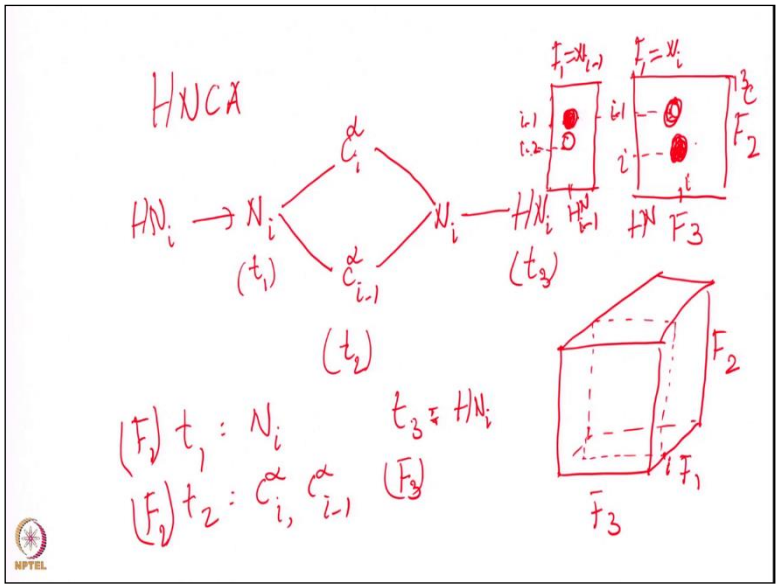
Cα and to this Cα, and from the Cα it comes back to this nitrogen from here and from here also it comes back to the nitrogen and it is transferred back to the amide proton and you detect it here.

Now, you label the nitrogens and you also label the Cαs and then you detect the amide protons as the amide proton chemical shifts in the third dimension. Therefore you have 3 dimensions here that is why it is labelled as HNCA. So, one dimension is the proton, another dimension is nitrogen, in the third dimension is Cα. So, here you see it correlates the 2 consecutive amino acid residues. If this is the residue of a particular amino acid and this is the NH of a particular amino acid,

so, this goes as NH-Cα-CO and this is the previous residue. So, we see correlations to the Cα of the same residue and also to the Cα of the previous residue at the N15 chemical shift of this particular residue. If I call this as i, this is (i – 1). So, I will have the correlation of the amide proton and the nitrogen-15 of the residue i to the Cα of i and (i – 1). So, this is how the magnetization transfer goes. Of course, you will have both peaks. Now this experiment provides a directionality here.

So, from this amide proton, you go to this nitrogen and then from here, you do not go to the Cα directly, but you go to the carbonyl and then from the carbonyl you go to the Cα and then from Cα you come back to the carbonyl and come back to the nitrogen, then come back to the amide proton to record the signal. So, therefore you label the amide protons and the nitrogen and the Cα. Therefore it provides a directionality. You will get only a correlation to the (i – 1) residue Cα.

(Refer Slide Time: 07:34)



Let us try and draw this picture in a little bit more explicit manner. We are going to talk about the HNCA. So, we start from the amide proton of residue i and we transfer the magnetization to the nitrogen of residue i. So, here what we do is, we put a time label here t1, which means we are generating an independent time domain. So, the magnetization is evolving on the nitrogen and this is the t1 time domain, where the time is incremented systematically.

So, this provides an independent time variable. So, now from here you transfer to Cα of i and Cα of (i – 1). Now you frequency label here also, what that means is the magnetization comes on the Cα, you evolve as a function of time and this time is also incremented systematically. So, this generates the second time variable. Now from this one, you come back to the nitrogen of the same residue i.

But you do not do anything here. Of course one can do experiments in 2 different ways either you can use this as a particular time increment or you can also use this one. Either way one can do these experiments. So, once you have this magnetization back onto the nitrogen, you transfer it back to the amide proton here of the residue i. So, this is the time variable t3.

So, what chemical shifts are present during the time variable t1? So, in the time variable t1, I have this nitrogen chemical shift of residue i, and in t2, I have the Cα chemical shifts of i and Cα of (i – 1). I have both these chemical shifts and then in the t3, I have the chemical shift of amide proton of i. So, therefore what sort of a spectrum I will get is a 3-dimensional spectrum here.

So, this will go to F3, this will be F1 in the frequency domain and this will be F2 in the frequency domain. Suppose I write here F3, this is the amide proton chemical shift and I write this as F1 and I write this as F2. So, let us say I have the N15 of residue i here, then what I do, I suppose, I take a cross-section here. So, I take this dotted line plane, what does this contain and the corresponding amide proton.

So, I will draw that particular plane indicated with the dotted line. Along this axis, I have the F2, along this axis I have the F3 and this is the chemical shift of residue HN of i. In the F2 dimension, I will have 2 chemical shifts in C13-dimension.

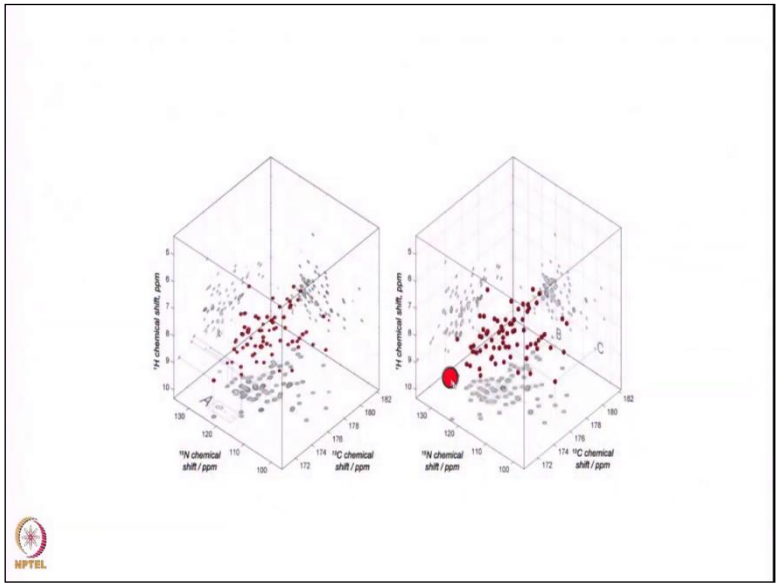
So, one of them is this is the C 13 of i and this is of let us say (i – 1). It could have been either way also or the other way around. It does not matter but one of them is i other one is (i – 1). This is at a particular F1 equal to Ni, N15 of residue i. So, what I have here, I have at the NH proton chemical shift amide proton chemical shift of residue i, I have the Cα of residue i and Cα of (i – 1).

Now suppose, I go to the Ni, chemical shift of residue (i – 1), so what I should see there? If I go to the plane of N(i – 1), that is at the F1 = N(i – 1), then at the N15 and the amide proton of (i – 1) again I should see 2 peaks and it should be the same Cα chemical shift. It should be here. These 2 should match right, they should be at the same place.

Then I will have another peak which will correspond to (i – 2). Let us say I draw that here this will be (i – 2) and this will be (i – 1). Now, this is the so-called self-peak and this is the sequential peak, in this plane this is the self-peak and this is the sequential peak. So, a priori you do not know which the N15 of (I – 1) is. So, therefore what one will have to do, one will have to scan through these N15 planes along the F1 dimension and find out where this peak appears exactly at the same place as this along the Cα chemical shift. You should be present at the same place. The amide proton chemical shift can be different of course because the amide proton chemical shift of (i – 1) and i can be different. Therefore when you search through the N15 planes, look for the same Cα chemical shifts which are present in the 2 planes. Then wherever this is present that will be the amide proton chemical shift of (i – 1) and that is what is indicated here.

So, you go from the amide proton to the nitrogen here, then you go to the Cα of i and Cα (i – 1), and then you go to the Ni, that is how this kind of a pattern will appear. Once we have this then you scan through the N15 planes, you will go to the amide proton of (i – 1), and then it will continue. So, you can keep drawing these kinds of planes. So, you will have.

(Refer Slide Time: 14:48)



So, if I draw those planes here one after the other. So, this one is F1 = Ni, this is F1 = N(i – 1), this F1 = N(i – 2), and this is F1 = N(i – 3) and so on. So, I will have here, let us have 2 peaks here one peak here as self, self is filled; and one is sequential with the open and then and this will be the HN of i.

Now I come here, I will get this is a chemical shift of HN(i – 1). Then I will also get here a sequential peak this will be (i – 2). So, then from here, I will go, let us say I get a peak here and this will be the HN of (i – 2) and there can be another sequential peak in the same plane and that will be of (i – 3). So, then from here, I will find HN(i – 3) and I may again find a sequential peak.

Now you notice here, the empty circle connects to the filled circle from one plane to the other plane. So, this is how we walk along the polypeptide chain sequentially connecting one residue to the next residue, only you will have to scan through the N15 planes. So, this is the strategy that is used. What happens in the HNCOCA experiment? We did not go to the residue i, we only went to the (i – 1).

Therefore in that case what will happen is, in the HNCOCA experiment, you will see only the sequential peak. So, for HNCOCA experiment, if I draw here peaks for the HNCA and if I do the HNCOCA experiment in the same plane I am only going to see this one. In fact, this is the strategy which is used to distinguish between which is the self-peak and which is the sequential peak.

And this is of HNi and this is of (i – 1) Cα and these are all Cα. Suppose there is confusion in this HNCA as to which one is the self and which is the sequential Cα peak, you make use of the HNCOCA experiment and you can get the identification of which is the sequential peak. So, that is how you distinguish and you can walk along the polypeptide chain.

Now these 2, therefore form a pair of experiments. So, let us go back and see what other experiments, we looked at here HNCA and HNCOCA, these are the pairs. For HNCO I may not draw want to draw this experiment in the same manner as I do earlier. So, therefore here you have from the amide proton you go to the nitrogen, you go to the carbonyl.

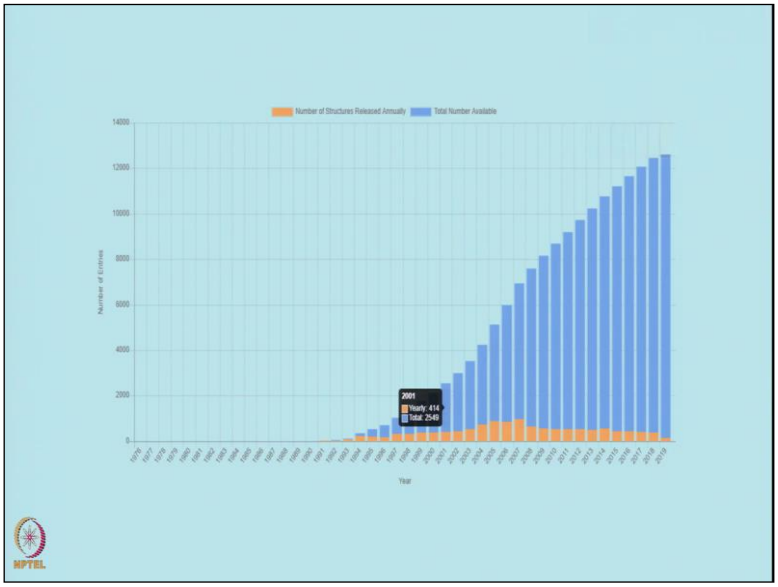
And you increment the carbonyl, increment the nitrogen, increment to the amide. Therefore in the 3-dimensional spectrum here you will have an amide proton chemical shift on one, nitrogen on the other, and the carbonyl of the previous residue on the third dimension. So, therefore this is a different experiment which will give you a carbonyl assignment. These ones will give you amide, nitrogen and Cα assignments.

But once you know the amide and the nitrogen assignments, you can also get the carbonyl assignment. So, in the same manner, you can go to this experiment, this is HNCA and this looks a little bit more drawn out because it goes much longer in the sequence. We can see here also the way it goes. So, you have from the amide proton to this here the nitrogen, from this to the Cα, also to the Cα.

Then it can also go to the Cβ here but you now label the Cβ, you are not labelling the Cα which means when it is on the Cβ, you systematically increment the chemical shift. However, often the Cα and the Cβ may not be separable and you may have sometimes both peaks present. So, you have this amide proton to this Cα to the Cβ you label this then come back here, come back here, then you from here you come back here, come back here and here.

Therefore you will get both these kinds of connections here. So, there are a whole lot of such experiments where you can record different kinds of chemical shifts along the different axis. So, this has allowed an enormous variety in the experimental sequences and you can calculate the structure of larger proteins.

(Refer Slide Time: 21:02)



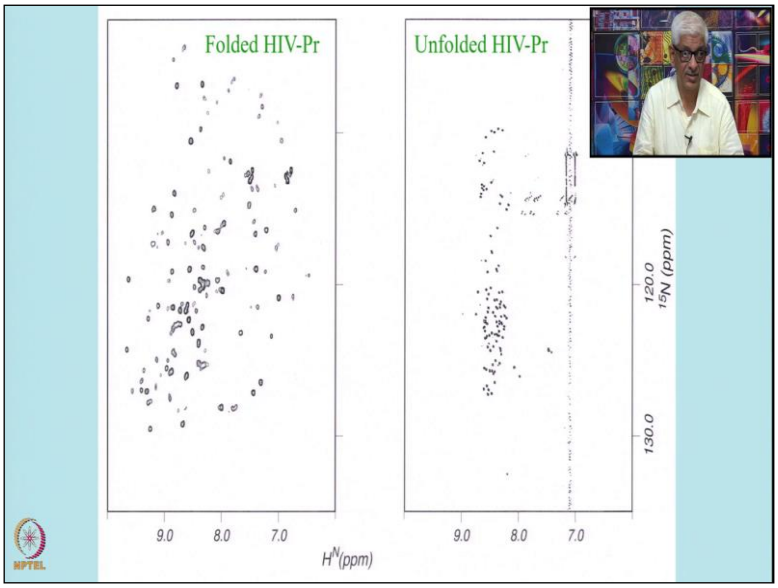
A 3-dimensional typical spectrum will look like this here and you have this amide protons on one axis, N15 on one axis, and C13 on the other axis. Now this is actually a stereo picture, If you have stereo glasses you can actually see this as one stereo picture where all the peaks are present here indicated, now the projections are also shown here.

Depending upon which plane, and which projection you take, you will get those correlations if you look at this projection, what is present here on one axis, you have the amide protons, other one is nitrogen. So, you will see like in the proton nitrogen HSQC spectrum. So, this axis is carbon, this axis is proton. So, this spectrum here is the proton C13 correlation spectrum.

Suppose I take this projection in this plane, what does it contain along this axis, you have the proton and this axis is the nitrogen and therefore this spectrum will be the proton-nitrogen 15 HSQC spectrum. So, these different projections you can take of such a 3-dimensional spectrum you will get those correlations. If you take the third projection here along this down there which is there. So, what is present along this, you have the C13 chemical shift and here is the nitrogen-15 chemical shift.

So, therefore this projection shows you the C13-N15 correlations. So, in one experiment you have all these possibilities of obtaining correlations of different types in a 2 dimensional way when you take the projections you will get this sort of data in your spectrum.

(Refer Slide Time: 22:48)

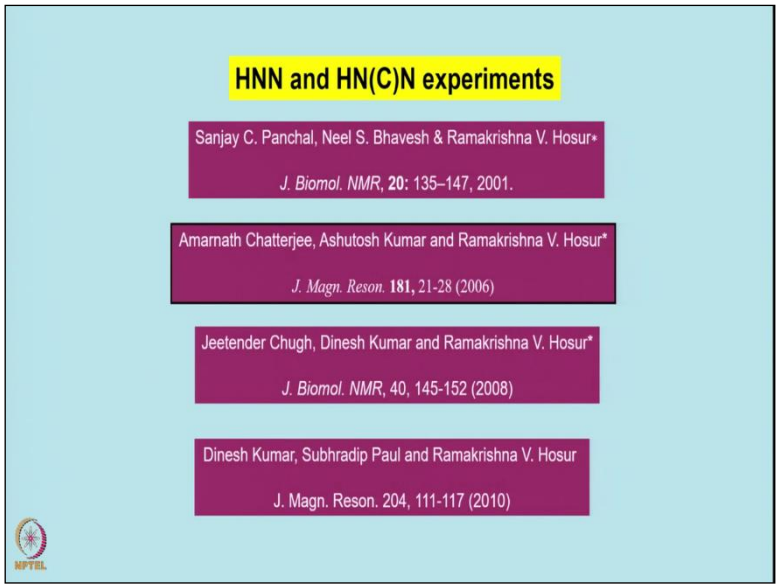


As a result of all of this, of course, you are able to handle larger and larger protein systems and you are also able to do a larger number of protein systems in a small amount of time. A few experiments are there and there is no question of ambiguities or overlaps of peaks because the dispersion is very good. You are using the chemical shift dispersion of both nitrogen-15 and carbon-13 and therefore a large number of protein structures have been determined.

And these have been deposited in the PDB data bank and this is a data which is of 2019. You can see how many protein structures have been deposited here. This initially was taking several years to determine one protein structure but with the development of these methods, this number of data deposition has increased quite rapidly and you can see how the progress has happened over the years.

And at this point, you see the total number is something like about 12000 protein structures up to 2019. So, this will continue to grow and now with the number of experimental sequences developing you get better and better structures. Now the emphasis is also getting towards larger protein systems and also complexes and the interactions because those ones will be separate issues and we can talk about that separately.

(Refer Slide Time: 24:08)



Now there are other kinds of protein systems, which pose challenges, and that is the so-called intrinsically unfolded proteins or it is also called IDPs, Intrinsically Disorder Proteins. These proteins pose the challenge because this is exemplified here. You can see this is a folded protein, HIV protease, and this protein if you denature it becomes an unfolded protein.

Then you see the HSQC spectrum looks like this. So, the peaks dispersion which was present here, so much dispersion which was present here has been lost, what is lost? The amide proton dispersion is lost. The N15 dispersion is still very good. Now when this happens actually the Cα dispersion also will get lost, when the amide proton dispersion is lost, it turns out that the Cα dispersion also becomes very poor.

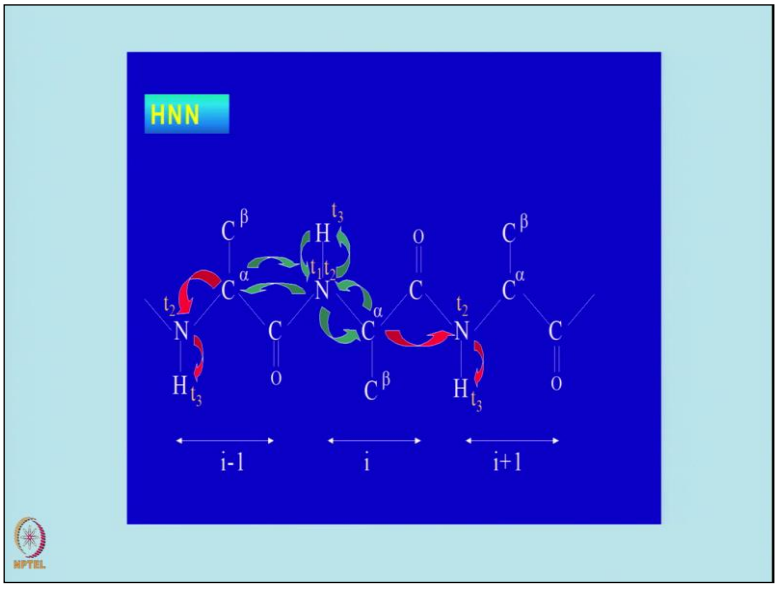
In that case, the HNCA-based experiments will have difficulties because there will be a lot of degeneracy in the Cα chemical shifts. Why that does not happen? Let me go back and show you here. So, what we showed here, suppose we have this sort of a chemical shift we are searching through the nitrogen-15 planes. Suppose I have this sort of a peak present here on more than one nitrogen plane, I identified a particular nitrogen chemical shift.

But suppose there are some 3 or 4, where the same peak appears, same Cα chemical shift peak appears, then you do not know which one to pick as N(i – 1). So, that is where if the Cα chemical shift dispersion disappears, you will have difficulty in identifying which one to choose and therefore you will have difficulty there. So, what one should do we have to devise some other methods where we make use of the nitrogen-15 chemical shift more than the Cα

chemical shift because the nitrogen chemical shift dispersion is very good even in this situation as well.

So, several people have put efforts in this direction and we also did some work in this direction and we published a series of papers and we will look at this series of experiments which we called as HNN-HNCN series of experiments.

(Refer Slide Time: 26:38)



So, how do these experiments work? Basically, it is very similar to the HNCA but it goes a step forward to label the nitrogen-15 more than the Cα. So, let us look here. So, let us say I start from this particular amide proton and I transfer the magnetization to this nitrogen as in the case of HNCA and I label this as t1. From this nitrogen, I go to the Cα here and also to the Cα. In the HNCA experiment, the Cα and the Cα were labelled as t2.

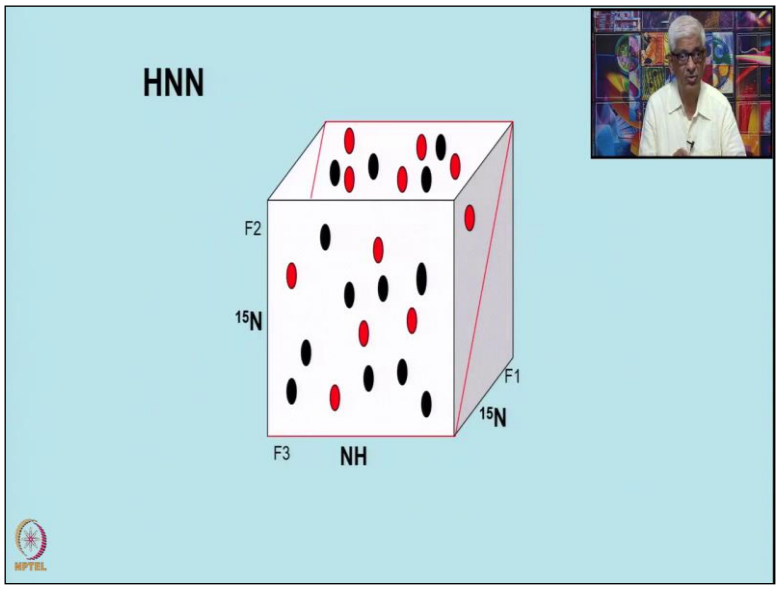
The time was incremented here but I will not do that. What I will do, instead of going labelling here Cα, I transfer the magnetization to the nitrogen here. Similarly from this Cα to this nitrogen, of course, part of it goes from this Cα to this nitrogen as well. So, when a transfer from the Cα to the nitrogen, what transfers will happen? The transfer will happen from here to here, it will also happen from here to here, then from here it will transfer here and also partly to this.

Now I am back on nitrogen. All these 3 nitrogens are present, now we label this. So, this becomes time increment t2. When I do this time increment t2, along the t2 dimension, I will again have the nitrogen, but this time the nitrogens of residue (i – 1), residue i, and also residue

(i + 1). These 3 nitrogens will get labelled during the t2 time period. From there I transfer from the nitrogen to the amide proton.

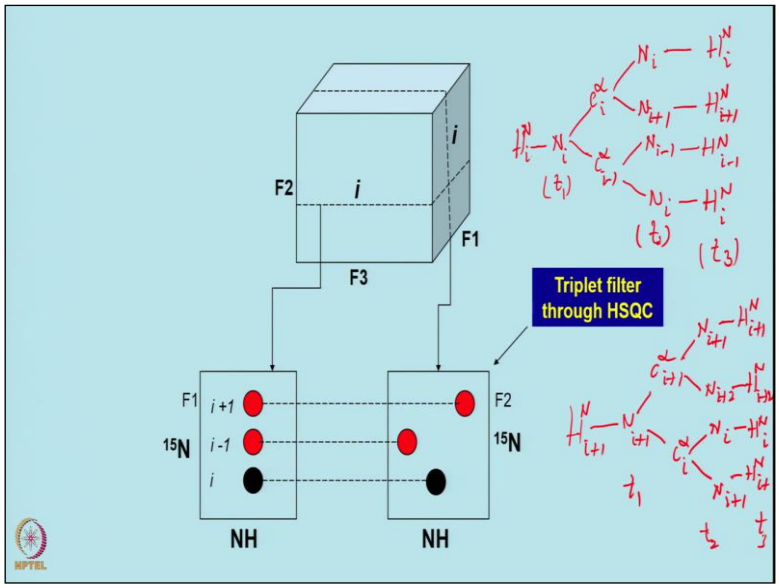
So it goes to this amide from here, goes to this amide from here, it goes to this amide from here. So, now I detect as a function of time t3 period, I have the amide proton chemical shifts. So, I have the amide proton chemical shifts of all the 3 residues of (i + 1), i, and (i – 1); all the 3 will get labelled.

(Refer Slide Time: 28:58)



So, the result will be. So, this is what will happen. So, I get a 3-dimensional spectrum where on the F3 dimension I will have NH chemical shifts, F1 I have also N15 chemical shifts and F2 Nl5 chemical shifts. So, what will be the details of this, how does the detail look like in this, that is shown here.

(Refer Slide Time: 29:20)



If I took this cross-section here at a particular F2 chemical shift. So, this is the particular F2 chemical shift I took this plane, what do I have there? Then that is my F1-F3 plane, this is my F1 axis, this is my F3 axis, and what I should get because starting from the particular amide proton of i, I had transferred to the nitrogen of (i – 1) and also to the (i + 1).

So, therefore at the particular F2 chemical shift of residue i, we can go back and see there, the F2 chemical shift of residue i, because this is F2 here, F2 here, and F2 here, and I started from t3 here. So, I will have a correlation to this nitrogen, this nitrogen, and this nitrogen, all the 3 will be present at the particular F2 chemical shift of residue i.

Similarly, go to the residue (i- 1) I will have different ones. So, let us go back there and see. So, if you look at the F2 chemical shift of residue i, I will have 3 peaks NH of residue i, this is the N15 of residue i, N15 of (i – 1), and N15 of (i + 1) as well. But if I now take the orthogonal cross-section, if I take this cross-section here that is at the F1 chemical shift of the residue i.

So, what will I have? I will have this at particular the chemical shift. So, this axis is F2 and this axis is F3. So, F3 is here, F2 is here, and this is at the particular F1 chemical shift. What will I have? I will have 3 peaks the residue self-peak i to its own N15, then I will have the (i – 1) NH to its own N15, (i + 1) NH to its own N15. So, this is how I will get these 3 peaks here. So, these 3 peaks will tell me that these are the 3 peaks from the entire HSQC spectrum. Therefore, we call this the triplet filter through the HSQC spectrum.

Let me try and draw this schematic here a little bit more for you to understand that. So, let us say I have here HN of i, I go to N of i, then I go to this side let us say I call it as t1 here, then I go from here to Cα of i and Cα of (i – 1). So, from this now I go N of i, then N of (i + 1), and from this, I go N of (i – 1) and N of i. Then I go from here HN of i and this will go to HN of (i + 1), this will go to HN of (i – 1) and this will go to HN of i.

Now this is my t3, this is my t2. Therefore you can see here HN of i appears at these 2 places. Where does it get contribution from, it gets contribution from this N I, that is coming from this pathway. So, you can write similarly coming from HN of (i – 1), HN of (i + 1) where does that thing come from? This is also showing you N (i + 1) and it also shows N (i – 1). HN of i is showing me to N i and is also coming to N i, but it is this pathway also generates this N(i + 1) and N(i – 1).

Suppose I started from HN(i + 1) here, then I will also get HN(i + 1) and HNi. I will show you that here. Suppose I want to start from HN(i + 1), so what will I get N of (i + 1) then I go to Cα of (i + 1) and Cα of i, then I go from here to N(i + 1), then N(i + 2), then from here what do I get? Ni and N(i + 1). So, you see now this one will give me HN(i + 1), this will give me HN(i + 2), this will give me HNi, and this will give me HN(i + 1).

Now you see here also I am getting HN of i, starting from this is my t1 here, this is my t2, and this is my t3. So, in this also I am getting HN of i, that is why I said here from this peaks here you can see at HN of i, I see i to i and (i + 1). Similarly, if I started from HN(i – 1) then also I will get HN of i. That is why I get these 3 peaks there. So, this is how this thing works and you will get in this particular plane all these 3 peaks in the same amide proton chemical shift.

Immediately you know which is the nitrogen-15 plane of residue (i – 1) and which is a nitrogen-15 plane of (i + 1). You do not have to scan through the nitrogen-15 planes as you have to do in the HNCA and if you look at the orthogonal cross-section you also get the NH chemical shifts of those 3 residues. So, this is how this experiment works and this simplifies your analysis much better there is no searching through the process and there is no degeneracy problem immediately you can identify the correlations. So, I think we can stop here.