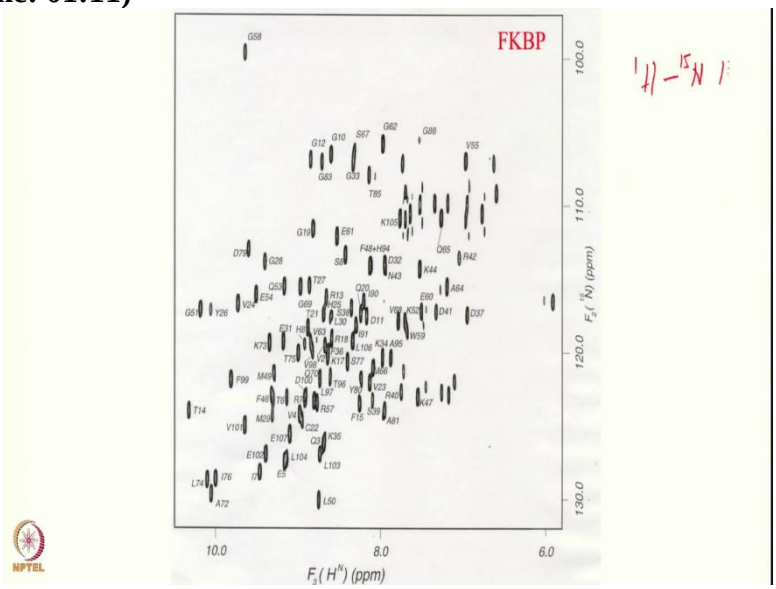
NMR spectroscopy for Structural Biology NS
Prof. Ashutosh Kumar and Prof. Ramkrishna Hosur
Department of Chemistry
Indian Institute of Technology - Bombay

Lecture: 22
Introduction to NOESY and HSQC 3

So, we have been discussing about multi-dimensional NMR methods. We discussed the different types of 2D correlation experiments and last time we also talked about the heteronuclear experiments and I want to show you some of those spectra and illustrate how these heteronuclear experiments are extremely useful. How we take it forward to 3 dimension 4 dimensional experiments as we also mentioned earlier.

But we will elaborate on that one as we go along different kinds of experiments are can be done. Now one of those which was most crucial experiment in this endeavour was the so called HSQC heteronuclears single quantum coordination experiment.

(Refer Slide Time: 01:11)



And that is HSQC and to this one what I am showing here is the so called proton N15 HSQC of a protein which is labelled as FKBP it is a small protein of about 90 residues and you can see here peaks coming from the individual amino acids as they are labelled here. You can see the various amino acids are labelled with their one letter code. So, for example there is a glycine here there is glycines here all these are glycine.

Then you have a serine here glycine (G) glycine valine and incidentally these figures which are coming here these are coming from the side chains side chains of glutamines and asparagines side chains also have the NH 2 groups CO NH 2 groups of the glutamines and inside the CO and H 2 group in the side chain and they also produce peaks here and they appear as pairs here. So, whereas all the other peaks which are appearing here.

These are from the backbone of the protein and you can get this one peak per residue I mean every amide proton has a connection to its own N15. Therefore you can see here such a good distribution of the peaks this is one of the very well behaved very well folded proteins. Now if you can imagine here how this experiment has been extremely useful see suppose you were to

take a projection here take a cross section at a particular NH position in the proton spectrum how would it look like?

Suppose I were to take the amide proton at this chemical shift and you see how many peaks are here 1, 2, 3, 4, 5, 6. So, many amide protons are present at this particular position this is what I illustrated to you last time that when so, many amides are present they will all come at the same line in your correlation spectrum of the TOSCY or the NOESY and things like that but now you see all of these amino acids have different ^{15}N chemical shifts.

And therefore they get separated here on this particular ^{15}N axis here. Similarly if you look here there are so, many amide protons at the same NH chemical shift and they have fortunately they have different and 15 chemical shifts. Similarly if you were to take the particular ^{15}N chemical shift somewhere for example here. So, then you will see so, many ^{15}N s at the same point.

So, you have here suppose I took this one this one this one this one this one this. So, many N15s at the same point N15 chemical should but then they have different amide proton chemical shifts therefore they get separated out quite distinctly in this 2 dimensional correlation spectrum amide proton N15 correlation spectrum that is why it is called as heteronuclear single correlation spectrum

Therefore since you are able to count here the number of peaks distinctly one for each residue the number of peaks will be almost equal to the number of residues in your protein barring prolines. The prolines do not have the amide protons therefore you will not show those cross peaks and therefore this experiment is called as the fingerprint of the protein with regard to its primary structure.

That is the amino acid sequence fingerprint of the protein with regard to the primary structure. You can just count the number of peaks here first of all you count the number of peaks to find out whether you got your protein properly or not or there something else is happening if there is aggregation going on in your protein. Then you will see that some number of peaks will not be the same some peaks do not appear.

Then this will also give an indication as to what is happening in your protein whether the protein is aggregating and if it is aggregating there will be only few peaks which will be seen and these will correspond to the flexible portions of the aggregate. Therefore this also you make use of to find out the status of the protein in your solution. And then of course you have to devise conditions how to break these aggregates and how to get single thing.

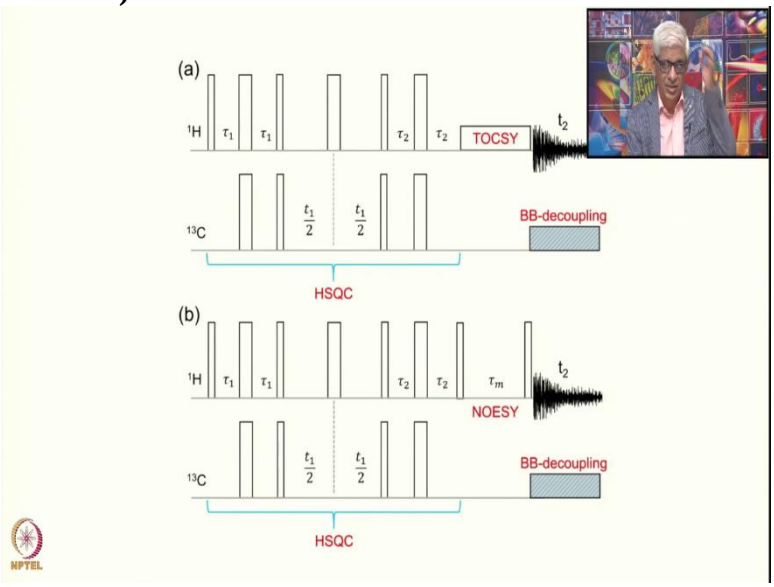
And this is what extremely is a useful technique and we will show you how this has been used to identify or determine the folding pathways in proteins or aggregation pathways in proteins this also has been extremely very well character useful procedure for this. Now of course what you need for this? You need a labelled protein. So, you need here a ^{15}N labelled protein. You cannot do this in at natural abundance.

And in fact this is how this was these experiments were not done earlier when the methods were not available for labelling the proteins ^{15}N labelled proteins are required for this and this is not possible because of the recombinant technology. You can produce this protein inside E.coli bacteria and you can put your gene of interest as kind of a construction construct and the

protein will be expressed in the cell you feed the bacteria with ammonium chloride which is ^{15}N label.

And then it will and no other source of nitrogen in that one then it will incorporate the amino acid the ^{15}N in every amino acid it synthesizes and then in every protein it synthesizes. Therefore you will naturally get ^{15}N label protein from the bacteria you will as a biological machinery to label this protein with ^{15}N and then of course you will have to devise methods of purifying your recombinant protein and then record the NMR spectra.

(Refer Slide Time: 07:08)



So, now having done that so, you as we also discussed earlier that we; can make use of this strategy and combine this with the other 2D experiments like the TOCSY and the NOESY. The HSQC spectra gave you only the amide proton chemical shifts it identified the amine proton chemical shifts and the N15 chemical shifts but the amino acids have other protons as well along the side chains and along the backbone the alpha protons the beta protons or the gamma protons and things like that.

So, therefore you should be able to identify those ones and also if you are going into the structural part then you must be able to identify which protons are close by. And this is what is done by these particular kinds of experiments where you have this HSQC part from here HSQC which were discussed earlier this is the HSQC part you combine this with the toxic spectrum here TOCSY this is a TOCSY mixing sequence.

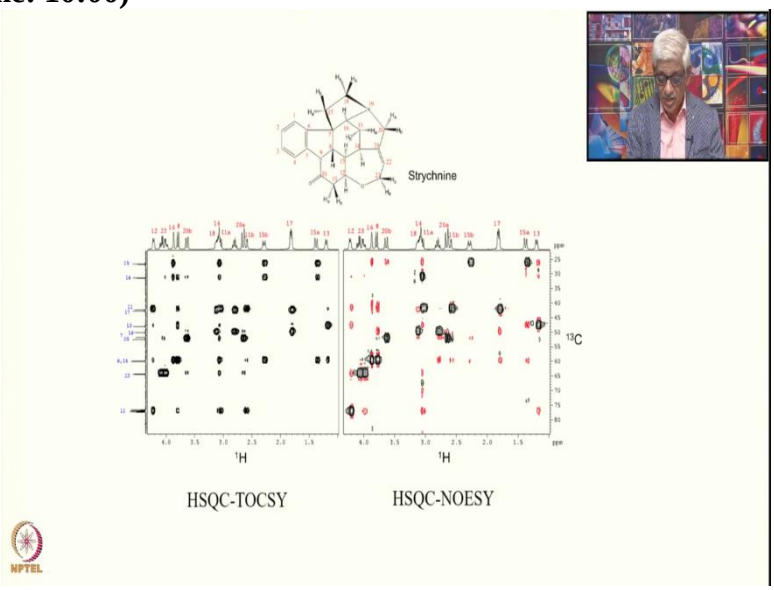
Until here it is the HSQC up till here then you make combined with the TOCSY part here and then you record the spectrum in t_2. Therefore what will appear along the t_1 axis the t_1 axis the information that is present is the ^{15}N or the ^{13}C here in the case ^{13}C has a pulse sequence is written but the same thing is true for the N5 as well so therefore you can this will be the carbon 13 chemical shifts or this ^{15}N chemical shifts in the case of proteins.

But in other molecules you can do carbon 13 chemical shifts here I will show some examples of this one then after this TOCSY you can record this spectrum what happens during the TOCSY period the magnetization which has come back on a particular starting proton will get relayed to other protons through the TOCSY mixing. So, this TOCSY mixing relays as we discussed before to other protons through the sequence.

Therefore at a particular carbon 13 you will see all these protons which are connected in the network similarly in this particular example until here you have then HSQC spectrum. So, the C13 proton correlation here and then after this, this portion which is added here this is the NOESY. So, you have the mixing time here the τ_m is the NOESY mixing time. So, during this period you have the transfer of information happening in the between the protons which are close by in space.

Or if there is a chemical exchange happening that also can be happening here. So, whichever process is happening you will see correlations for those particular protons therefore this has a structural input. So, while this is a useful for identifying the spin systems the assignments this has the structural input it also has input from the point of view of assignments because the near neighbour interactions will provide you the assignments and it will also provide you the structural information as in your depending upon the molecule what you have.

(Refer Slide Time: 10:00)



Let us see here the example here. So, here is an example of a small molecule this is strychnine this is some kind of a drug I guess. So, this is a molecule which has been investigated in detail this is a very complex molecule. So, you have. So, many protons here all the protons are numbered here see you are 21 22. So, many protons are there is carbons these are all carbons. So, many carbons are there and is starting from one here 1 2 3 4 5 6.

So, many carbons are present in this molecule and there are correspondingly so, many different protons. Now to determine the structure of this to assign this individual spectrum it is extremely complex. So, we look at this chemical shift this is the one dimensional spectrum of this molecule. So, you have. So, many peaks here of course this is the same here as well and this side is the HSQC TOCSY spectrum this is the HSQC NOESY spectrum.

There is a big difference between these 2 and you can see the ones which are the dark ones here focus on the dark ones here the red black peaks these are the HSQC peaks alone. These are just HSQC peaks. These correspond you can count here the number of peaks this will correspond to the number of carbons you have in your molecule. So, each one of them here corresponds to one particular carbon.

Each peak corresponds to one particular carbon and you have here the number of pace which are equivalent to the number of peaks in carbons in your molecule. Of course if you have a

quaternary carbon that will not show up here because wherever there is a proton attached to the carbon only that carbon will show up here. So, you can have as many carbons identified there. Now what you do at the same carbon positions you see this TOCSY.

So, you have the whole range of protons which are attached. So, from here to here there are so, many protons which are connected to that. So, different carbons there and from the carbons you have a relay to the protons which are attached therefore you can identify the protons which are connected to these carbons. So, you see here these are ones which are connected in the coupling in the network can I you can look at these peaks here.

So, which are the ones which are connected the numbers are indicated here. So, this assignment has come as the result of analysis of these spectra. So, you have the carbons and you have the connected protons to them and you go to the next carbon which is connected to that then you walk through the spectrum to identify the entire network of carbons which are connected to each other.

So, that is how you get a sequential walk or a walk through the spectrum through the TOCSY spectrum to identify the various carbons in your molecule. What about this? This peak here you see you have the HSQC spectrums which are black and the NOESY peaks which are present here these are appearing in red the red means they have negative peaks these are negative peaks. Why are these negative? NOESY, because this is a small molecule.

You remember in a small molecule which is less than few hundred molecular weights or few hundred molecular weights the tumbling is extremely rapid in those cases the NOE is positive. If the NOE is positive in the 2 dimensional spectrum it will appear as a negative peak whereas the diagonal or not the angle the self peaks appear positive and the NOE peaks will appear as negative.

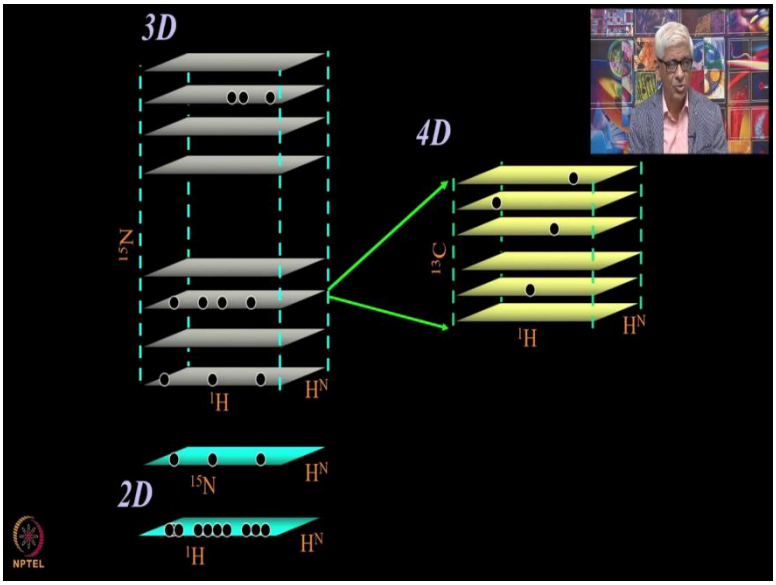
Therefore you can clearly distinguish which are the NOE peaks and which are the self peaks here that is why I said all these black ones are the self peaks of the individual carbon proton correlations it will tell you how many carbons are there in a particular carbon chemical shift and from those carbons you have a NOEs to the protons. So, what does that tell you that means these protons are having correlation through the distance dipolar distance.

And therefore you can identify using this we can figure out which protons are close by here because the NOE can happen between the protons and protons. We are talking about proton-proton NOE we are not talking about carbon proton NOE. We are talking about proton-proton NOE. So, let me let me go back to this experiment here you see the magnetization at this point is on proton magnetization is on proton.

So, therefore we are talking about the proton-proton NOE. So, therefore when I see here so, the NOEs which are happening are between the protons if I have carbon here it is attached to a proton and it shows NOE to another proton and it is attached to a different carbon then you will see that sort of a correlation here we can walk through the walk through the carbon chain here from the proton-proton correlations appear in the particular carbon chemical.

This is just to illustrate how these experiments are useful for arriving at the assignments on the one hand and the structure of the molecule on the other hand.

(Refer Slide Time: 15:02)



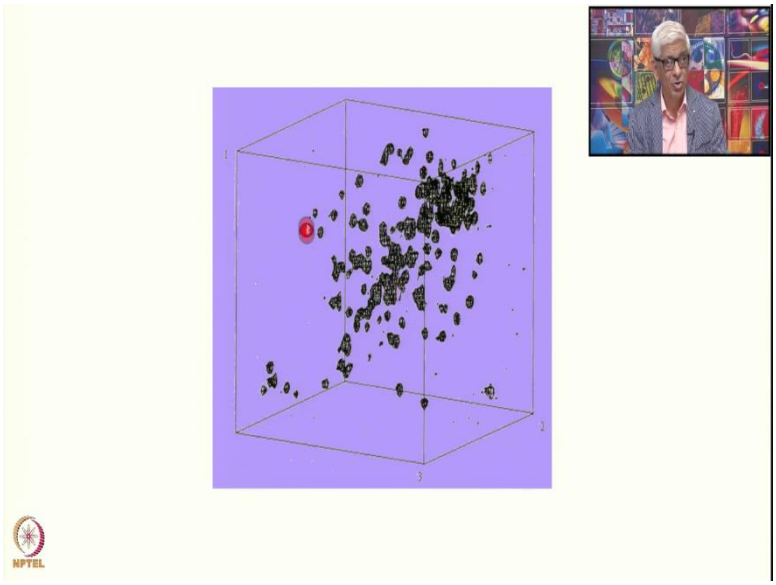
So, now you can see you appreciate this better the how we can make use of three dimensional experiment. Now we talked about the two dimensional experiments we said that a particular I explained to you earlier at a particular amount proton chemical shift there are so, many amino acids here. So, therefore you are seeing. So, many chemical shifts. So, many peaks here this may be a TOSCY spectrum this may be NOE spectrum whatever.

So, you have. So, many peaks here at a particular proton chemical shift. Now if you look at the proton N15 correlation spectrum you are seeing three ^{15}Ns therefore there are at least three or certainly three because there are only three peaks there. So, you have three peaks which are present here and. Now you do what you do is instead of making a just a 2 dimensional spectrum HSQC spectrum HSQC TOCSY or something like that.

Instead of spreading them along the proton axis extending the proton axis here like that we actually pull this spectrum along this axis. Here in the 2 dimensional spectrum what we did we would actually pull this in this is HSQC NOESY to include the proton chemical shifts more here. But now we will restrict that to ^{15}N the amount chemical shifts only but we will take the third axis to pull these peaks onto the different dimension the ^{15}N dimension here.

So, therefore all these peaks will. Now get distributed into the different planes in the 3 dimensional spectrum. Now if you need further resolution you can make use of the carbon 13 here then of course you will go into the 4D experiment and this will be quite a substantial thing to do. Now if you want to do carbon 13 then you will have to have a carbon 13 labelling as well. This also can be done using the bacteria and use carbon 13 labelled glucose and you will get a ^{13}C and 15 level protein.

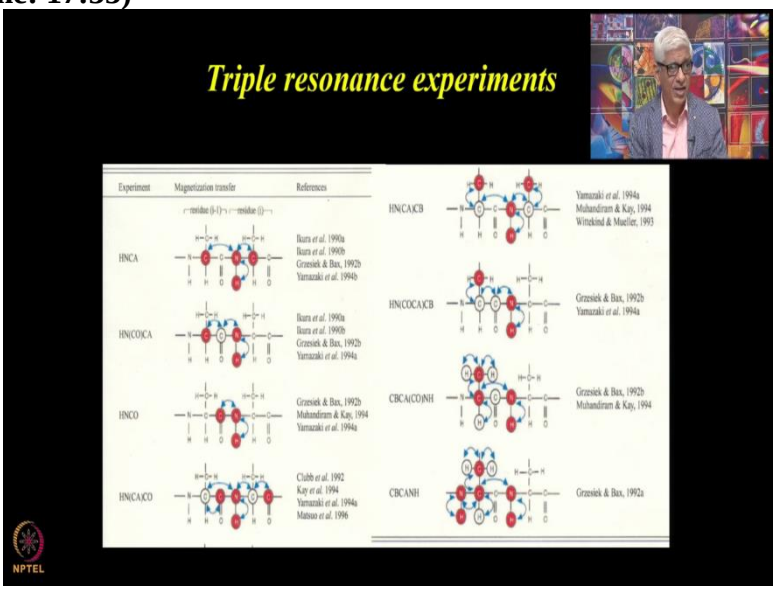
(Refer Slide Time: 16:51)



Here is a typical example of a 3 dimensional spectrum here axis are not labelled with anything but I already showed you yesterday that in the previous class then you can have whatever nuclei you want here you can have you can have 2 proton axis or one carbon axis two proton axis or one hydrogen axis or one proton axis one nitrogen axis and one carbon axis. So, you can have all 3 different kinds of combinations there.

So, now the peaks are distributed in a volume in a cuboid rather than on a plane therefore you take cross sections at various places you will be able to figure out what sort of a spectrum you will get what is the information present in your spectrum.

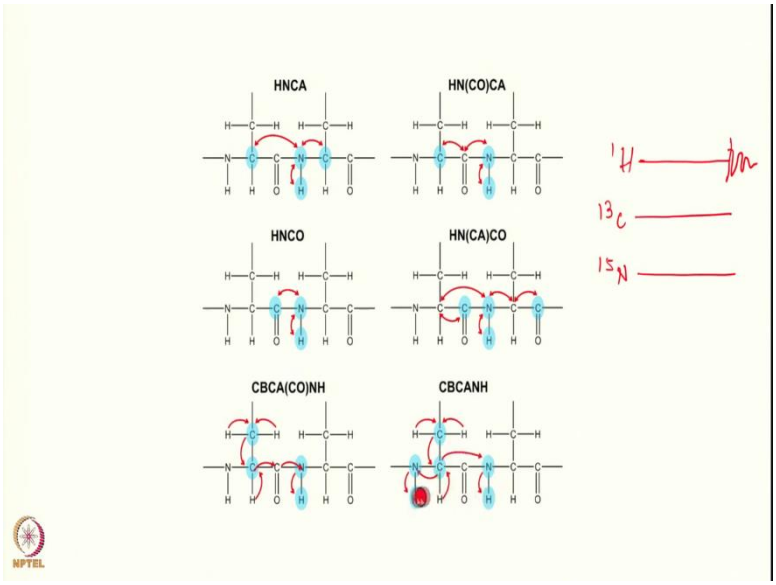
(Refer Slide Time: 17:33)



Now we go to the next stage, we go to the next stage what are called as triple resonance experiments. So, far we are talking about 2 dimensional experiments or 3 dimensional experiments where there were only 2 nuclei one was the proton other one was the nitrogen. Now you can have 3 different nuclei also right as I mentioned you can have proton nitrogen and carbon what does this require this requires.

Different kind of treatment you need to have 3 different channels you need to have 3 different channels I will write this on the next slide here.

(Refer Slide Time: 18:11)



This is the same slide given in the white background. So, that you only some experiments have been selected here. So, you need to have in your spectrometer you need to have a proton channel a carbon channel and a nitrogen channel. So, you need to be able to apply pulses along the 3 different channels. So, you apply various pulses there apply various pulses and how it will and obviously finally you actually collect the data only on the proton channel.

So, in between the magnetizations flows to through different ones which I will explain to you we take one by one and see how this works. Let us look at this particular experiment which is called as the HNCA what does this represent this experiment represents that along one axis you have the proton this is proton and along the second axis you have the carbon these are the Cα carbons and on the third axis you have the nitrogen.

The red ones are the nuclei which you are going to see in a 3 dimensional experiment that is why these are called as triple valent experiments because you are making use of proton nitrogen and carbon 3 different channels. And these experiments were described way back in about 1990s by different authors here. So, I have listed those references here for quick reference seeing through that.

Now this experiment starts from the proton here and the magnetization is transferred you see how this is manipulated. You manipulate the transfer pathway of the pulse sequence you go from here to here, here to here or here to here and then come back here come back here and come back to the amide proton and detect it this is exactly the strategy what we used in the HSQC TOCSY or HSQC NOESY.

We went from the proton to the nitrogen and we came back to the proton to observe it. Now we are doing 3 different things here we go from the proton to the nitrogen and you also have the frequency label there and from the nitrogen you go to the carbon to the both the carbon C alphas you also have a frequency label there. So, you generate another independent time variable and after that you come back here come back here and come back to this and you generate a detect the proton signal.

So, now this is called as the HNCA. Now the next experiment here is the HNCOC there are so, many different kinds of experiments here and the red labels the ones which are nuclear marked

red are the ones where you have the frequencies labelled. So, in this case you have the proton the nitrogen and this C alpha. Notice we do not go to this Cα we go to this Cα.

How do we do it because we go we in this case we went directly from the nitrogen to the Cα we transferred magnetization from the nitrogen to the Cα here we do not do it we transfer from the nitrogen to this carbonyl we transfer to this carbon and from this carbon we transfer to this Cα. Therefore there is no question of going to this Cα at all this provides a kind of a directionality in our assignment procedure.

From this amide proton we go to this nitrogen from this nitrogen we go to this carbonyl from this carbon we go to the Cα then we come back to this carbonyl come back to the nitrogen and then to the proton for detection we detect the signal on the on the proton. So, here therefore in the 3 axis what are the 3 axis we will have again proton nitrogen and Cα . But this is Cα only now.

Now we do another trick here this experiment you have the amide proton the nitrogen and the carbonyl you label this carbonyl come back to the nitrogen and detect this on proton therefore we label these three. Here I have proton nitrogen carbonyl these are the one peaks which will appear in your spectrum. So, likewise you have various possibilities of transfer pathways and this is manipulated by using different kinds of pulse sequences.

We are not going to go into that one but we see we know the principles how to transfer magnetization from one nucleus to another nuclei another nucleus and we simply apply those principles here. So, what we do in this experiment. So, we transfer from the amide proton to the nitrogen from this nitrogen we transfer to the Cα here this Cα as well as the Cα this is the same as this, this is the same as that.

Now what we do from the Cα from the Cα we transfer partly to this nitrogen and partly to this carbonyl partly to this carbonyl and. Now we put the red label on this carbonyl we are not putting the red label on the Cα as we did in this particular case. So, we use this is the mediator here as the Cα come to the carbonyl and then we from the carbonyl we go back to the Cα then back to the nitrogen same here C carbonyl to the Cα.

Then to the nitrogen then here therefore these ones have different kind of information as you will see. And similarly you can also do that on the side chains of the individual amino acids this is called as HNCACB or correspond on the HN(COCA)BC on these are various experiments which are designed to walk through the amino acids different carbons in the same amino acid residue.

So, in this case we actually start from these beta protons here from the beta protons we start we come to this Cα here and then from the Cα we transfer to the nitrogen from the nitrogen we transfer to the amide proton and we detect the signals here. Therefore what are the signals which detect here we detect Cβ the nitrogen and the amide protons. We finally detect this thing always notice we detect the signal on the mind we do not detect the signal anywhere else.

Because the amount protons are very distinct and we can easily identify those ones there. In this particular case you again start from the Cβ protons here from the beta protons you go to the Cβ carbon from the bond you go to the Cα here and then from the Cα you go to the carbonyl and unlike here from this Cα we went to the nitrogen directly but here we do not do that.

We go from the Cα to the carbonyl from the carbonyl we go to the nitrogen from the nitrogen we collect onto the amide proton and therefore this is called as and what is written inside here is what is written inside the bracket is these are the ones which are on the pathway these are there on the pathway but we actually do not detect those signals. Detecting means we have a frequency label.

And the t_1 increment should correspond to that particular chemical shift and that is what determines which nucleus you are going to detect. So, yeah and then the CBCA(CO)NH so, these were the HN protons. Now here what we do we start from the Cβ once again from here to here. So, we have from the Cβ you come here from the Cα also you start here you also start from the Cβ.

So, both the things will come here and from this you go to the nitrogen and from the nitrogen you go to the amine proton you and you detect the signal here. So, in this in these 2 experiments you can also start from the amide eventually in fact start from the amide proton amide proton to the nitrogen, nitrogen to the carbonyl, carbonyl to the Cα from Cα you go to the Cβ.

And then from the Cβ you go back to the Cα then to this carbonyl and to the nitrogen and to the amide proton and detect it here. So, you can do that either way. So, therefore this kind of these are called as out and back sequences. So, different kinds of strategies one can use to achieve the your magnetization transfers. All this the way you the way you start and where you end will determine the way you design your pulse sequence.

From whether you want to start from the alpha proton or the beta proton or from the NH protons and where you would finally come back and detect it that is but ultimately you always want to detect on the amide protons. In this particular experiment you see 2 amide protons here. So, you have this NH proton as well as this NH proton. So, you have this ones coming once again it goes through the same pathway go from here to here, here to here and here and they also can go from here to here.

Or you can go from Cβ to this then from here you can go here then. So, these are all out and back pathway and from here you start from the Cα you go to this nitrogen as well. So, you have transferred to the NH proton at this point.

So, this is I explain that in a little bit more explicit manner in the few of these experiments here. So, you have HNCA, HN(CO)CA, HNCO these ones provide a crucial assignment strategy here in this CBCA(CO)NH experiment this is much more clearly indicated in this slide than in the previous one. So, in this experiment you see you start from the Hβ from the Hβs does you go to this Cβ from the Cβ you go to the Cα then from the Cα you go to the carbonyl.

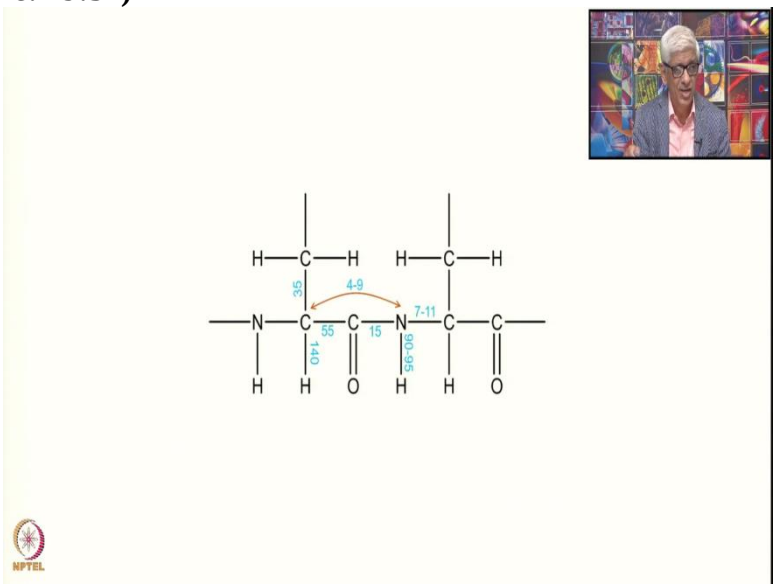
And in this you can also the some portion of the magnetization will come from the H alpha also because the Hα and Hβ which are rather difficult to distinguish with regard to the chemical shift the pulses will apply to both. So, therefore it will come from here to here as well and then from here to here as well both the things will happen then you go to the NH amide ^{15}N and then to the NH proton and for detection here.

And in this particular case the CBCANH you have the starting from here Hβ or Hα you come down to the Cα is it a 2 step process here Hβ to Cβ and then to the Cα and from here also you go to the Cα from Cα you go to the nitrogen and from this α you also go to this nitrogen and you collect the signal as amide protons chemical shifts at this point.

Therefore these will establish correlation what correlation do they establish they will establish correlations between within the 5 amstrong residue up to the Cβ carbons but you see this experiment will also establish a sequential correlation. Because it has the 2 amide protons involved here of 2 consecutive residues this is one residue here and this is the next residue here.

So, therefore you will establish correlations not only within the same amino acid residue but from one residue to another amino acid residue.

(Refer Slide Time: 29:34)



Now these experiments are dependent on the transfer pathways and how is the transfer occur the transfer occurs through the J coupling and the J couplings are these are all one bond J couplings one bond J couplings and these numbers are indicated here see the one bond advantage is on one bond or 2 one the advantage is that these are not dependent on the confirmation of your protein and these are always the same.

Therefore there is no question of missing a correlation here in this kind of experiments here. So, you have this various coupling concepts that are indicated here these are the carbon-carbon coupling 35 hertz carbon proton coupling is 140 hertz this coupling curve this Cα CO coupling is 55 hertz and carbon nitrogen coupling this one bond is 15 hertz this 2 bond coupling nitrogen Cα is 4 to 9 hertz.

And in one bond nitrogen Cα coupling is 7 to 11 hertz nitrogen to proton coupling is 90 to 95 hertz all of these do not vary depending upon the amino acid do not vary too much on the conformation depending on the conformation therefore you will not miss any correlation unlike in the case of NOESY where you can miss a correlation if the distance is longer if kind of a motion is such that you are not able to see a NOE.

But here you will not miss any peak therefore this is an extremely useful strategy to obtain sequence specific resonance assignments and for large proteins reasonably large proteins which you could not handle through the NOESY spectra because these are using making use of the amide ^{15}N and carbon chemical shifts. All the 3 chemical shifts are used and therefore the dispersion is enhanced and you will not miss any assignment.

So, so we will stop here and we will explain a little bit more about the sum of these pulse sequences to give an idea of how the transfer happens in these pulse sequences as an illustration. We are not going to discuss every experiment in this but I will take some 2 illustrative experiments to show you how the manipulation is done through the pulse sequence I think we will stop here.