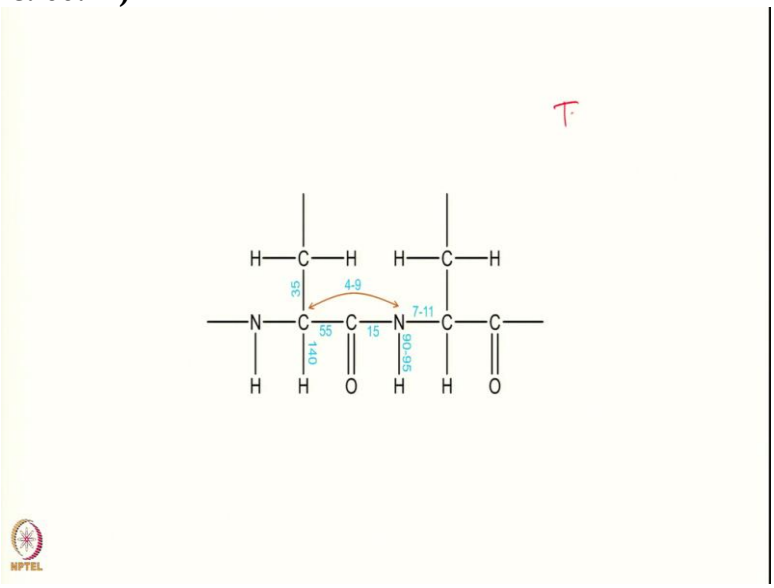
NMR spectroscopy for Structural Biology NS
Prof. Ashutosh Kumar and Prof. Ramkrishna Hosur
Department of Chemistry
Indian Institute of Technology - Bombay

Lecture: 23
Introduction to NOESY and HSQC 4

(Refer Slide Time: 00:21)



So, you can see that the triple resonance experiments are basically driven by application to proteins experiments driven by for proteins. In fact this has been an extremely useful feedback system that multi-dimensional NMR experiments developed but the proteins actually drove them very strongly. So, desire to obtain structural information on proteins has been their motivation for developing new and new pulse sequences and you will see.

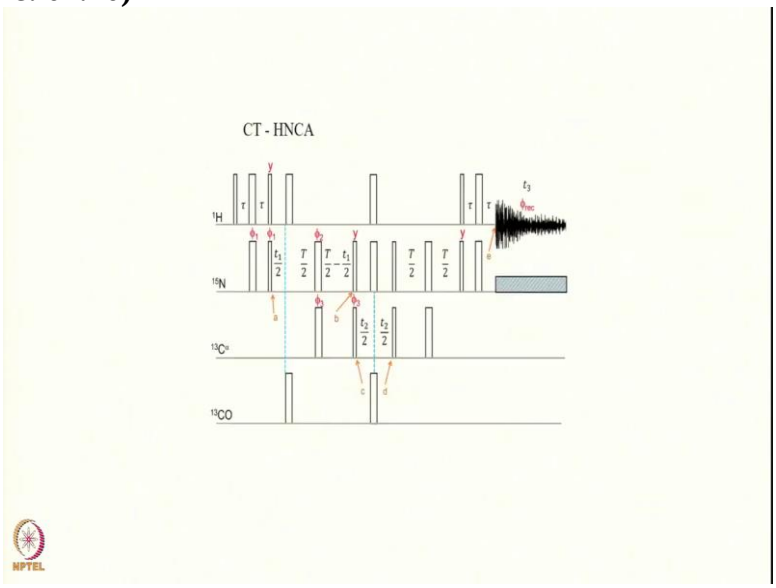
Therefore all these triplets resonance and experiments which we will describe are primarily directed towards proteins and therefore we will be concentrating on that kind of a thing structural biology primarily has developed as a result of application of NMR in structural biology has primarily been driven by application in proteins also of course nucleic acids and other systems. But by and large methods development has occurred because of applications to proteins.

So, here I show this slide again to you for the indicating the strategy here we have here a dipeptide NH, Cα, CO, NH, Cα, CO and the transfer of magnetization happens through this polypeptide chain make either through the backbone alone or through the side chains as well accordingly we have different kinds of pulse sequences here. And all these are based on transfers based on J correlations and therefore the J coupling constants are important here.

These different J couplings are indicated here these are all one bond couplings except this which is a 2 bond coupling and even the 2 bond coupling also does not vary too much as you can see and by and large it is about 4 to 9 hertz but mostly it is in a random call situation they will be like 6 to 7 hertz and certain structural situations you may have 4 hertz and somewhere it is 9 hertz and things like that.

But mostly it is one bond couplings and therefore there is no question of missing those peaks in your in 3 dimensional experiments. And how does this work I will illustrate this to you how the transfer happens in a particular amino in the protein in the pulse sequence.

(Refer Slide Time: 02:40)



And we will illustrate this is one taking one particular example which is called as the constant time HNCA I described to you the HNCA this is the another version of the same one the magnetization transfer follows in this particular path. So, it starts from the proton as you can see here and this is the amide proton this is here the amide proton which labelled here. The amide proton is starts from here and then you transfer from this amide proton using the INEPT sequence.

So, you have the initial INEPT sequence here. So, the magnetization flows from here and comes to nitrogen at this point. It starts here and comes to nitrogen at this point up soon after that part and then during the next period from here to here it flows in the nitrogen on the nitrogen path remains on nitrogen and then it is transferred to Cα. So, this is the pathway. So, the magnetization flows in this manner from the amide proton it goes to the nitrogen from the nitrogen it goes to this Cα.

And now you see in the amount when it is in the amine portion you have this $\frac{t_1}{2}$ increment the t_1 increment is present here therefore and if this total time from here to here it is t total time from here to here is t and that is why it is called as constant time but in the same one this t_1 is incorporated from here to here it is t therefore it is called as constant time. During that time the J evolution happens and during the t_1 period is the chemical shift evolution.

Which chemical shift that is nitrogen chemical shift because the you have the t_1 increment and the magnetization is on the nitrogen. Therefore during the t_1 period we have the nitrogen chemical shift. And now from this it is from the nitrogen it is transferred to the carbonyl to the C alpha and then over here you see $\frac{t_2}{2}$, $\frac{T}{2}$ and $\frac{T}{2}$ this is t_2. So, total is t_2 this is again in increment this is incremental time period.

And therefore what is the chemical shift labelling here if there is an incremental time period it means there is a chemical shift labelling that frequency is getting labelled with a particular evolution during that evolution that particular frequency is getting labelled. So, therefore

during the t_2 period I have this Cα getting labelled in the t_1 period ^{15}N is getting labelled and during the t_2 period Cα is getting labelled.

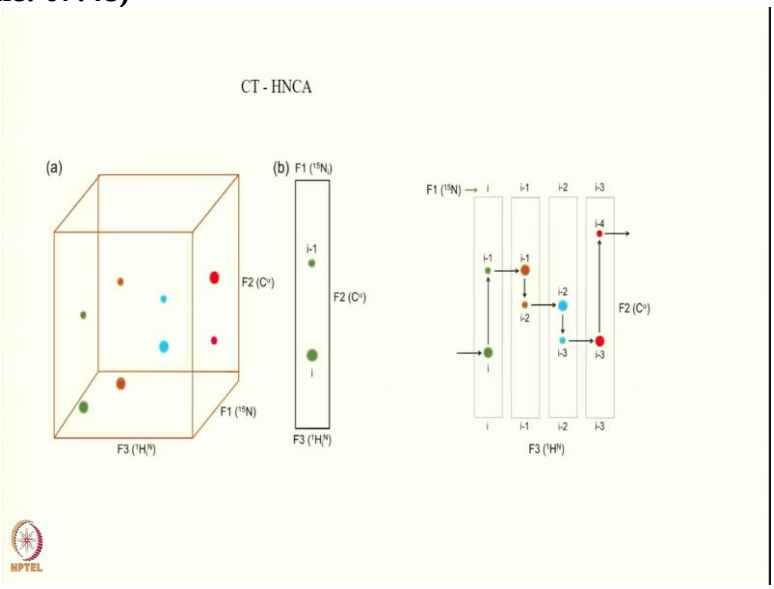
Now from here it goes flows here and then it goes to nitrogen again it goes to the nitrogen again and during the nitrogen there is no labelling here. So, this is the constant time again this is the same constant time as here T total period from here to here is T and during this time the whatever certain refocusing things like that will happen and it is on the nitrogen the magnetization is on the nitrogen.

And then it flows through the nitrogen until here and then it goes back to proton it goes back to proton here. And then during this period there is a refocusing of the amide proton chemical shift which is antiphase there and you produce at this point e in phase proton magnetization. So, that they can detect here and therefore what is in the t_1 period t_1 period I have the nitrogen chemical shift and t_2 period I have the ^{13}C see α chemical shift and during the t_3 period I have the amide proton chemical shift.

Because what I am going to detect here is coming from the nitrogen from the ^{15}N it is coming to the proton and where does it come it comes to the proton which is attached to it only therefore it is on the amide proton. And the amine proton is refocused here it is the antiphase it comes to in phase and then you detect this during the t_3 period. Therefore what the after 3D Fourier transformation what do you expect to get?

Now this is let us say if I call this as F_3 F_1 and F_2 what will be present in F_1, F1 will be nitrogen 15 and therefore this will go to F_1 this will go to F_2 and this will go to F_3. So, this is the way it goes of course we apply certain pulses on the carbonyl channel etcetera for the purpose of decoupling etcetera we may not want to go into those details here. But this is the way the magnetization flows and therefore this is the kind of a schematic spectrum you will get.

(Refer Slide Time: 07:48)



Now let us look at this spectrum how does it look this is what I indicated to you here and on this here you have the schematic spectrum of a particular sequence whatever randomly some 4 amino acids are taken here. So, how many amino acids are taken here there is one green there is one here red and then the cyan and then of course there is another red here this is brown. Now each one of them has a strong peak here and then also a peak at this point each one of them has that.

So, therefore it is indicating as I mentioned to you earlier you generate the self peak as well as the sequential peak of residue i - 1. So, if you let me go back there and show you what we had there HNCA. So, if you look at this. So, if you look at this we transfer we transferred from the amide proton to the nitrogen and nitrogen we transferred to both the Cα this α as well as this Cα.

So, this is the self residue, i and residue i - 1 therefore the both the things will have to reflect in your spectrum and that is what is happening here. So, you see you have here the 2 peaks at each amide proton chemical shift 2 peaks one belongs to its own C alpha other one is to the neighbouring C alpha which is i -1 residue same here. Therefore if you took a cross section at a particular then fifteen chemical shift and these ones are appearing at a different ^{15}N chemical shifts this is through the depth you can go through the ^{15}N here depth.

So, these ones are different depth these 4 ones are present at different depths. So, if I take a cross section at any particular ^{15}N in any particular ^{15}N this is the F_1 axis I have 2 peaks this one of them is the residue i other one is the residue i - 1 these ones are slightly different intensities by and large they may not have also but sometimes you do have slightly different intensities because this coupling constant is slightly larger than this coupling constant.

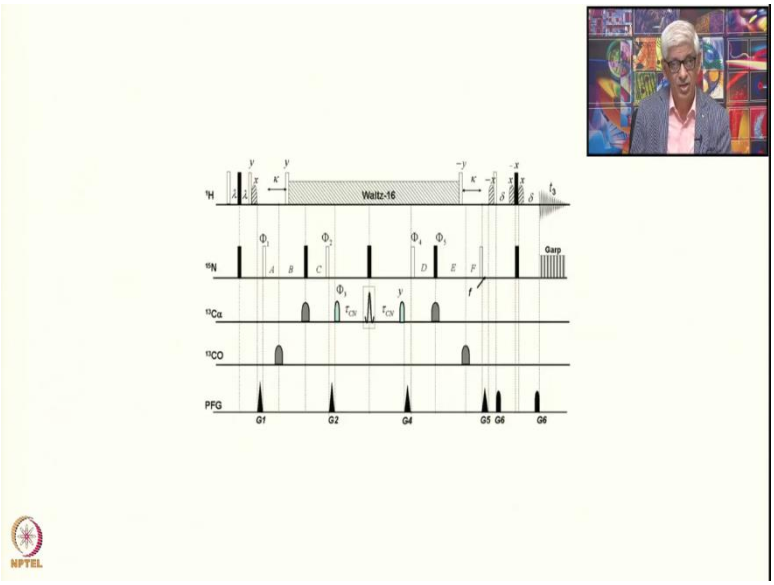
Because this is dependent on the two bond coupling whereas this is dependent on the one bond coupling as we saw this coupling constant is of the order of 7 to 9 hertz whereas this 7 to 11 hertz this coupling constant is of the order of 4 to 9 hertz. So, depending upon what it is sometimes you can have different intensities in fact it is helpful if it has different intensity it is helpful to figure out that which one is the self and which one is sequential.

Once we have got this. Now this is this sequential which is i – 1, residue. Now obviously at some other nitrogen chemical shift this will become the self peak when you move through the ^{15}N plane at this particular C alpha this peak will become the self peak because it has to have a particular self peak also. So, therefore you search through the ^{15}N plane here then you will find that at the particular chemical shift then of the ^{15}N you find exactly at the same Cα chemical shift you all you find a strong peak here.

This strong peak therefore is the self peak of residue i - 1 and then correspondingly you have another peak which is the weaker peak which is of i - 2. So, again you search through the nitrogen 15 planes then you will find a peak which corresponds to the same Cα chemical shift but at a different nitrogen amide proton chemical shift and also the different ^{15}N chemical shift. So, therefore you will see a strong peak here and a weak peak here.

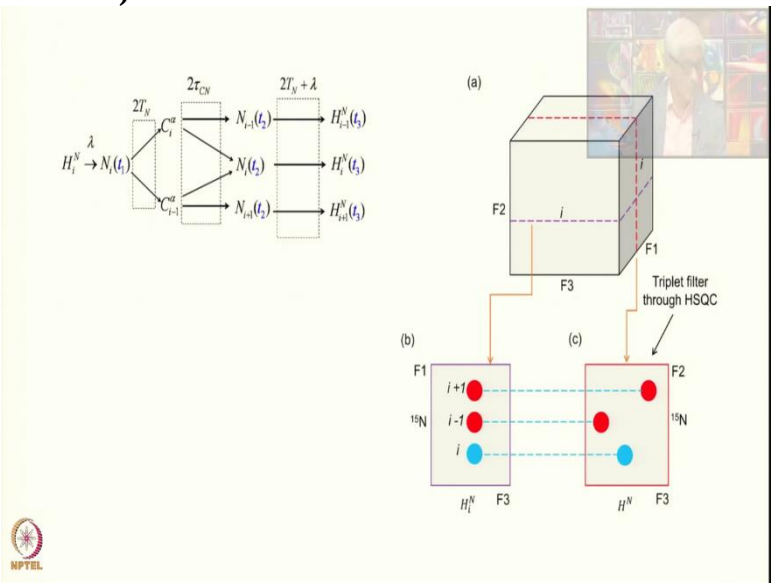
So, you can continue like that. So, therefore you see you can walk through the polypeptide chain by going through the different ^{15}N planes scan through the ^{15}N planes identify where the Cαs are and you can start connecting the peaks in this manner. Once you have the C alphas identified like this obviously you can also identify the other protons through the TOCSY or you can also use the other ones where you can go to the Cβs and so on.

(Refer Slide Time: 11:55)



So, this is an experiment which is extremely useful. Now here the separation will depend upon how good your Cα chemical shift is and there sometimes the Cα chemical shifts may not be too good and in that situation one has to use some other strategy and here is an experiment which is called as HNN. So, here what you do is.

(Refer Slide Time: 12:22)



I will explain to you the first the strategy how the magnetization transfer happens in this in this experiment the magnetization goes through starts with the ^{15}N amide proton starts with the amide proton and is transferred to the nitrogen of the same residue and it is frequency labelled with the t_1 there is a t_1 increment there. Now from this t_1 you transfer to the Cαs of the 2 residues i and i - 1 just as you did in the case of H and C a from N_i you transfer to the Cα of i and also to the Cα of i - 1.

Now in the earlier case you labelled this Cα but here we do not label the Cα we do not make it as a t_2 part what we do is we transfer this further to the next nitrogens there from this C alpha during a particular time period here called $2\tau_C N$ you transfer again to the nitrogen of i - 1 partly to the nitrogen of i - 1 and partly to the nitrogen of i itself.

So, this is N_i - 1 and this is N_i + 1 here and this will be i + 1. So, here again this will be i - 1 because the transfer happens from the N_i to the next residue an i of Cα i from the Cα it can go

to the i + 1. and Cα i - 1 its i + 1 will be i. So, therefore it can go to i here and part of it remains as i - 1 here. So, this is the way the transfer happens and after this you have got the proton you have the magnetization on the nitrogens of 3 residues i – 1 and i + 1.

And this is then transferred to the corresponding amide protons this will be i + 1 this will be i and this will be i - 1. So, these ones are detected during the t_3 period. And now if you do a Fourier transformation 3D Fourier transformation what you get here you get here a 3 dimensional spectrum F_3 axis has the amide protons F_1 has nitrogen and F_2 also has nitrogen. Now if we were to take a cross section if you take a cross section through the through this 3D spectrum.

How does it look like suppose I take a F_1 cross section at a particular F_2 chemical shift at a particular F_2 chemical shift I take this plane and show it here. So, this is the F_1 F_3 plane at a particular F_2 chemical shift chemical shift that is I have here particularly $F_2 = N_i$ and this is the N_i (H_i^N) of the F_2 then I will see 3 peaks these 3 peaks i, i - 1 and i + 1 all the 3 we see here.

And correspondingly if I were to take a F_1 cross section that means I take cross section through this here through this plane at this particular chemical shift then I will get here the F_2 F_3 cross section. The F_2 F_3 cross section has the same 3 peaks but at the respective amide proton chemical shifts see these are 3 different chemical shifts of the individual 3 residues these 3 peaks.

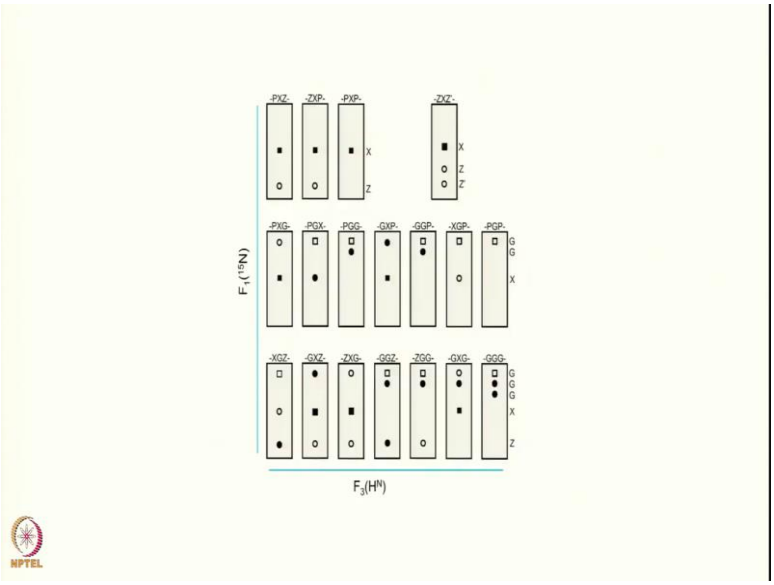
So, what do you see of this? This is like a triplet filter through the HSQC what do I mean by that. So, suppose I have a HSQC spectrum this is ^{15}N and this is HN you have lots of peaks here your many peaks for the 2D spectrum. But now out of these 3 consecutive residues whichever are the 3 consecutive residues and you will figure filter them out in this particular plane you will filter that in a particular plane.

So, therefore it is called as the at the particular F_2 F_1 chemical shift the particular ^{15}N chemical shift you get 3 peaks which are neighbours the self peaks i - 1 peak and the i + 1 peak all the 3 peaks you are seeing in this. Therefore this is called as a triplet filter through the HSQC spectrum and this is extremely useful because you immediately know which plane you have to go to identify the next residue.

You do not have to scan through the ^{15}N planes as in the case of HNCA to figure out where the Cα becomes a self peaks. Here you do not worry about the Cα first and the ^{15}N dispersion is always very good compared to the Cα chemical shifts in most proteins. And particularly in the case of disorder proteins intrinsically disorder proteins or loop segments in proteins you have more chemical shift dispersion for the Cα.

But the ^{15}N dispersion is always good therefore you can see this kind of peak separation in every case.

(Refer Slide Time: 17:56)



And another important thing you have to notice and that is I have used different colours here see I have used one particular colour for the self peaks the residue of the same residue and I have one particular colour for the two others I have two different colours because they have negative sign if this is positive sign this is the negative sign and vice versa whichever one you want to choose that way.

So, they have opposite signs and that also shows up here and this is extremely useful in a generalized sense but another important feature of this is this sign pattern what we are having depends upon what is the sequence in the triplet what kind of a sequence do you have in the triplet. Is there a glycine in the triplet sequence or not. Now if there is a glycine in the triplet sequence these colour combinations will change.

And that is an extremely useful factor because you you will be able to figure out where you are along the polypeptide chain as you are walking along the polypeptide chain if you hit the glycine you hit a different kind of a peak pattern therefore that will appear as a checkpoint. And that is demonstrated here in this particular slide what we are of course if there is a proline if there is a proline it will not produce a peak.

Therefore you will have only 2 peaks not 3 peaks because there is no amide proton on the proline but whereas in a generalized sense you will have the self peak which is positive and the 2 sequential peaks will have negative sign I represent this as Z, X Z', X is the central peak which is the self residue and Z and Z' are the i - 1 and i + 1 residue. These ones have negative sign as indicated in the previous one if this is positive.

The square indicates that it is the self peak or the of the same residue and the circle indicates it is a sequential peaks. Now similarly if I have here a proline with a glycine with a glycine there. Now the glycine is a very distinct chemical shift and with regard to the N15 and those ones will appear at the top in the N15 chemical shift. Therefore this will produce a distinct pattern you can immediately identify that this is the glycine.

Now if there is a glycine in the middle of a sequence if the glycine is here at the end then of course the the X residue will show a peak here and the glycine sequential peak will appear here if the PGX that means the glycine in the middle then you will have the square which is on the

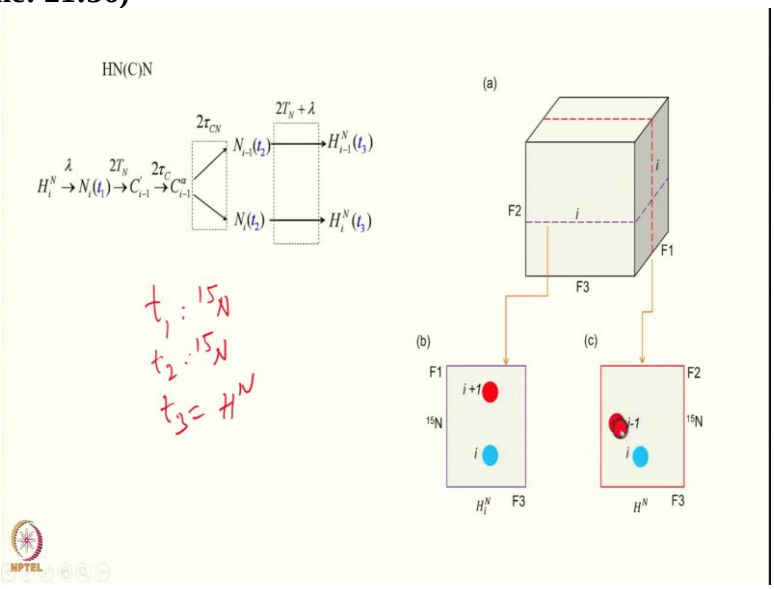
top but this will have a negative sign. See unlike this X which is a positive sign the glycine self peak will have a negative sign.

And you see everywhere where there is a glycine in the middle it has a negative sign and this is an extremely useful factor because you will immediately figure out that where you are along the polypeptide chain. Now these are the others where there is no proline and each one of them has 3 peaks each one of these peaks. If there is a glycine in the middle then that self peak is negative and then you have a positive and a negative combination here.

And if the glycine is at the i - 1 position then the self peak is positive and you will have a negative and positive combination negative and opposite combinations here. Therefore you see this peak patterns which you are seeing is extremely important in the walking along the polypeptide chain. You immediately identify are you doing the right thing or not you go from one residue to another residue.

And if you hit a glycine in the middle then you must get this sort of peak patterns positive and negative combinations and that will be extremely useful for rapid assignments of your polypeptide chain.

(Refer Slide Time: 21:36)



Now there is another complementary experiment of this which is called as HNCN. Now this experiment the magnetization transfer pathway goes like this. So, here you start from the again the amide proton and you go to the nitrogen 15 here. From the nitrogen 15 you go to the carbonyl you do not go to the Cα you go to the carbonyl of residue i - 1 from this carbonyl of i - 1 you go to the Cα of i - 1 during this period and then from the Cα of i - 1 where do you go you go to the nitrogen of i - 1 and nitrogen of i.

And here you type have frequency label you call this as t_2 and the t_1 is here the t_1 therefore along the ^{15}N axis you have the t_1 and here you also have the t_2. So, therefore then you have here and the knight from the amide from the nitrogen you go to the amide proton and you have here the resistance therefore what you have here and also in the same case in the previous one also.

So, t_1 is ^{15}N, t_2 is also ^{15}N and t_3 is amide this was all the also the case in the in the previous one maybe I should also write that here. So, here also t_1 is ^{15}N t_2 is also ^{15}N and t_3 is amide. So,

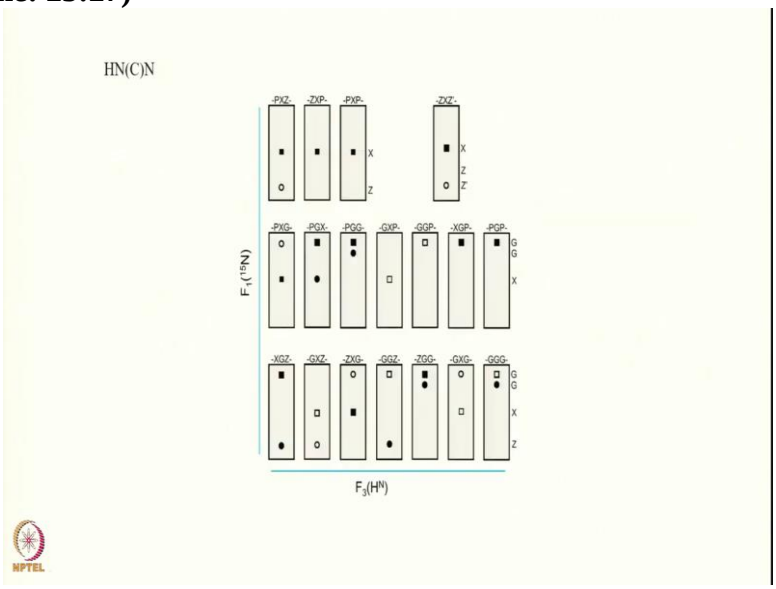
therefore you will have 3 dimensional spectra in this manner. So, now if I take 3 dimensional spectrum of this what is its content. So, let us look here you take the what are the correlations we will get?

See I am not going to the three residues there I only have two. So, if I take a cross section here through the at a particular F_2 at a particular F_2 position that means I have the F_1 F_3 cross section F_1 F_3 cross section has 2 peaks one particular to this own residue and then to the i - 1 start from the H_i^N I am seeing to i - 1 and i correct. So, so here H_i^N, i residue to i and i – 1, i and i + 1.

This will give you i and i + 1. Because you have to write down similarly for the H i of i + 1 also H_i^N where it will come you have to combine this you have to write such transfer pathways for H_i^N, H_{i-1}^N and H_{i+1}^N when you do that you will see that the particular H_i^N i you will see 2 peaks one corresponding to the i of the ^{15}N of i and ^{15}N of i + 1 you will see two peaks in the F_1 F_3 plane.

Whereas if you look at the F_2 F_3 plane F_2 F_3 plane that is at a particular F_1 which is the I that is this F_1, i what are you seeing you are seeing 2 nitrogen which are i and i - 1 therefore and they will appear at the respective H^N chemical shifts and that is what is shown here. At the respective H^N chemical shifts i - 1 and I you're seeing 2 different things. Now this is like a doublet filter I see here 2 peaks which correspond to the i and i - 1 residues along the polypeptide chain.

(Refer Slide Time: 25:27)



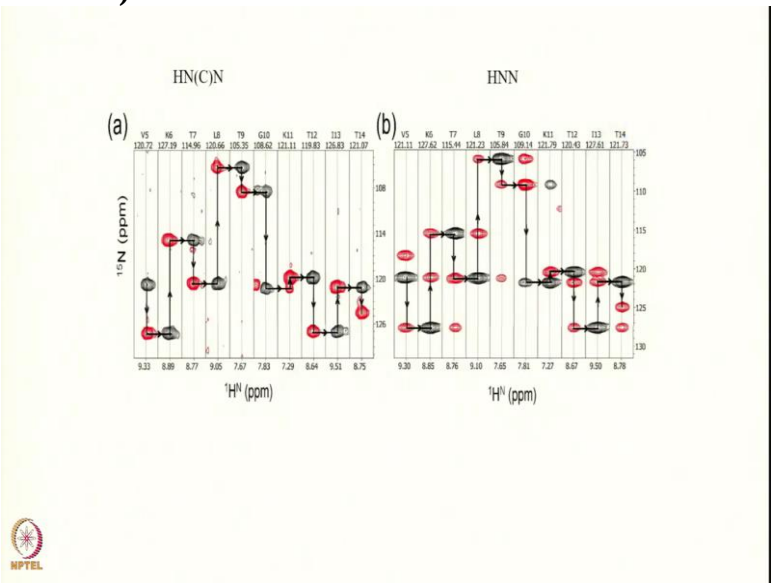
Now here again you will see interesting peak patterns. Once again you will see different kinds of positive and negative combinations and again once there is a proline of course you will see only 2 peaks here. But you will see 2 peaks in every case here but the patterns of positive and negative combinations are quite distinct compared to the previous one. Therefore this will allow you to identify which are the sequential peaks which are the self peaks and which direction you are going.

The important thing is this will provide you a direction in the earlier case where it was not easy to figure out which is i - 1 and which is i + 1. But now you see you have only i - 1 coming in this plane and in this play you are carrying only i + 1. So, therefore it provides the directionality for your walking through the polypeptide chain and that is an important observation and this will have significant value in your sequential resonance assignment.

This positive-positive here very unique this positive-positive for PGG sequence and for XGZ sequence also we see positive positive here. If you have glycine in the middle and you have X and Z again you are positive positive but if glycine is in the Ni - 1 positive you have negative-negative this is very interesting. So, if you have ZZZ then you have positive positive on the top here. So, if you have this situation negative-negative.

So, these are extremely important combinations of peak signs and which will you can cross check therefore these appear as check points in your walk through the polypeptide chain.

(Refer Slide Time: 27:16)



So, this shows experimental illustration of this spectra application here a particular protein is probably FKBP or Ubiquitin I do not remember but you can see here. So, sequential correlations how they are appending positive negative combinations see this is 5, 5. So, this is shown on the stop the sequence is given here 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, for the same 3 sequence here through the HNN spectrum how you are working and now we are doing it here for the HNC.

And so, you can do the sequential work through both the spectra whether you want to do it through HNCN or HNNN you can do that. So, you see if you go from 5 to here see the appear sign then you go from here to here here to here here to here your same positive negative combinations. But now as soon as you come here to the G10 see as soon as you come to the G10 you have to have two positive peaks there see it is TGK sequence TGK, G is in the middle and that must give you positive in the positive-positive combinations and this is what we saw here.

So, see this is, this, this portion XGZ. So, this g is in the middle and X and Z can be any other residue and as soon as you reach there you must get a positive-positive peak and in fact you see at this point you see a positive-positive peak. And the immediate next one should be negative-negative because that one is GKT that is GXZ. So, therefore you will have here negative-negative combinations.

And then after that again the G is gone. So, you have positive-negative positive - negative positive - negative and go on things like that. The same feature will show up here in this particular case as well. So, you have here in H^N as well you go through the sequential walk

from 5 to 6, 6 to 7 and 7 to 8 and so on. So, forth you can walk through the polypeptide chain along there and notice once more here.

So, as soon as you come here you see 2 negative peaks and this is the G10 just as you had here 2 positive peaks and you really get 2 negative peaks and the positive peak here. And the immediate next one will be 2 positive peaks and one negative peak and this is the guarantee that you are not made any mistake in your sequential work all through the polypeptide chain. So, this is therefore extremely useful for obtaining rapid assignments in proteins.

Regardless of the size especially in intrinsically disorder proteins where there are flexible domains and is very difficult to obtain assignments without this sort of a dispersion. So, N15 dispersion is always good. So, you can get resonance assignments in a rapid manner and in an unambiguous manner. This therefore experiments have been very successful in analyzing unfolding pathways or folding pathways of proteins or aggregation pathways or proteins and these strategies have been used.

And this is possibly we will discuss later when you actually discuss more of the applications of these methods and to draw more information about the biology of that one the structural biology of those one. Here our focus is on discussing the methods and how the methods are useful to derive the information and how they have been developed to circumvent some problems of peak dispersion or difficulties with regard to the assignments.

Or with regard to the structures in the proteins where there are unfolded regions in the folded regions together you will have difficulties. So, the methods development has been driven by such kind of problems and the applications of these we will also discuss more as we go into particular kinds of questions with regard to the folding unfolding aggregation etcetera. So, so I think we can stop here.