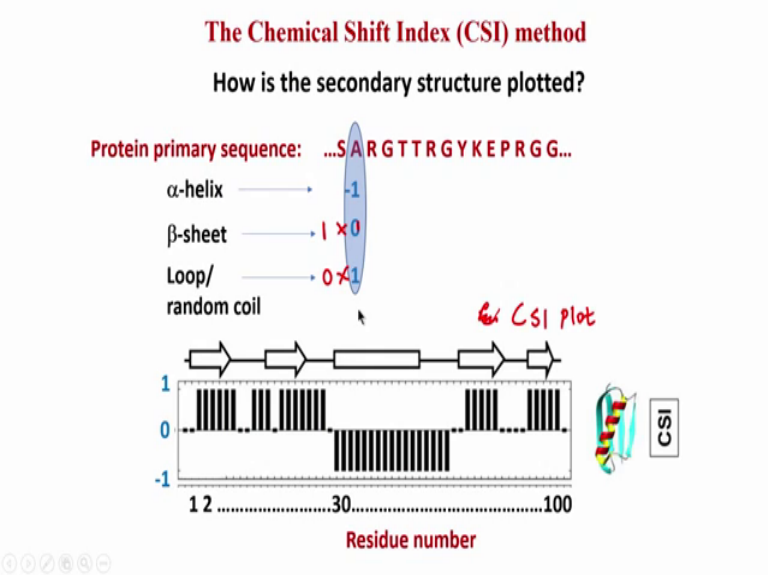
Multidimensional NMR Spectroscopy for Structural Studies of Biomolecules
Prof. Hanudatta S. Atreya
Department of Chemistry
Indian Institute of Science, Bangalore

Lecture – 32
Determination of protein secondary structure from NMR data: J-coupling based method

So, having calculated the chemical shift index as we saw in the last class, now the question is how do we show it pictorially because in the previous class, we saw we are calculating it for one amino acid at a time, but in naturally we want to see whether the entire protein what is the overall secondary structures present in the protein.

(Refer Slide Time: 00:47)



So, this is where we start plotting the CSI graph or the CSI plots so, how do we do that let us see here. So, let us say this is your protein primary sequence ok. So, this is how it goes from one end to the other end. So, what we do is if our the three chemical shift that we saw the three different criteria for alpha helix that is for H alpha should be less than point minus 0.1, beta C alpha should be greater than 0.7, CO should be greater than 0.5 so on, then we call it as alpha helix.

So, if all the three chemical shifts or 2 out of 3 agree then we say that that particular amino acid is an alpha helix or not. So, if that particular amino acid for example, serine here become is considered as alpha helix we will assign a number minus 1 to serine, just

a simple number its a arbitrary minus 1 is just a number assigned or given to S this particular S if it belongs to alpha helix. If this S according to this CSI approach was found to be in a beta sheet confirmation we will make it 1. So, let me correct this is a mistake so, this should be 1 and this should be 0 ok.

So, if this serine is found in a beta sheet confirmation based on this approach which we saw in the last class then that particular serine will be just given a random an arbitrary number 1, then similarly loop if it belongs to a loop region based again on this chemical shift values we will assign it a value 0. So, what we are doing is simply assigning every amino acid here either 1, 0, minus 1 based on whether it is in alpha helix beta sheet or not which was calculated using CSI method.

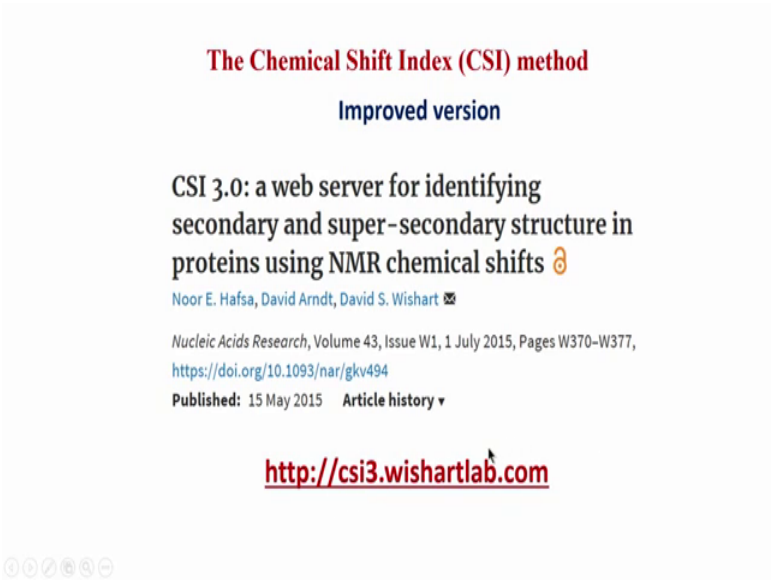
So, let me show you another example suppose I shift to the next amino acid then here again I will assign for alanine whether it is alpha helix beta sheet or not based on that I will assign 1 number to or these numbers to this amino acid. So, like this for every amino acid I will assign these numbers and then we will plot it like this. So, this is the plot called the CSI plot ok. So, let me write it here, this is called CSI plot. So, this is a very standard plot which is shown in many papers whenever you publish the secondary or a structure of a protein by NMR this is typically shown.

So, you can see here now in this graph this is for some protein like this which is a ubiquitin, you can see that as we go from one amino acid to others. So, let us say we start from those this here on the x-axis is the residue number residue number means starting from 1 going up to the last amino acid. So, if you see here the first two amino acid were found to be in a random coil confirmation so, it was 0 according to this here. Then, the amino acid next third onwards continuously they were found to be in beta strand, beta sheet. So, they were given 1 1 1 for everybody like so, this is shown like this in the pictured form.

Then again there was a loop a short loop random coil then again beta sheet, beta strand a short and beta strand and so on. So, look at this; look at this stretch here there was a long stretch where all their amino acids were having negative means secondary structure was helix alpha helix so, they all of them were given the number minus 1. So, there is a long alpha helix which is shown here this helix here on the right side you can see and then again your random coil beta sheet random coil and so on.

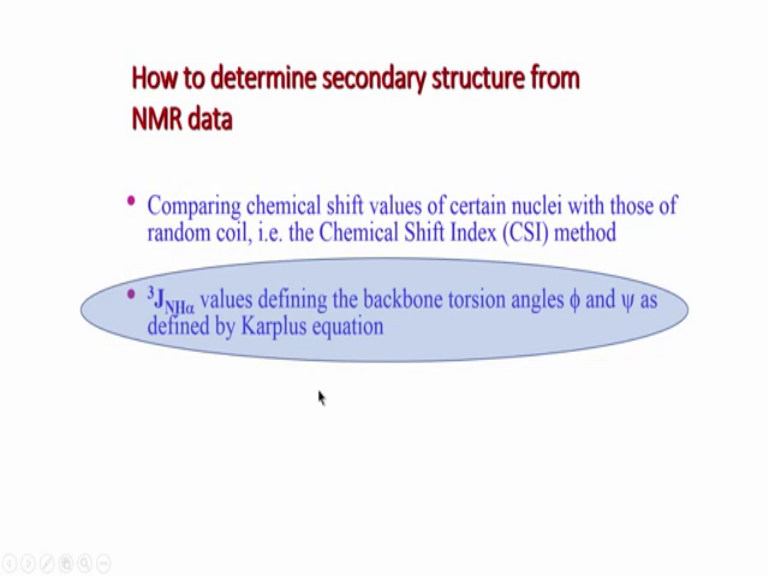
So, this is how we typically plot the CSI for a given protein to show or depict what are the different secondary structure elements in that protein. So, this is basically the CSI method which we typically do so, we will now move on to the next.

(Refer Slide Time: 05:14)



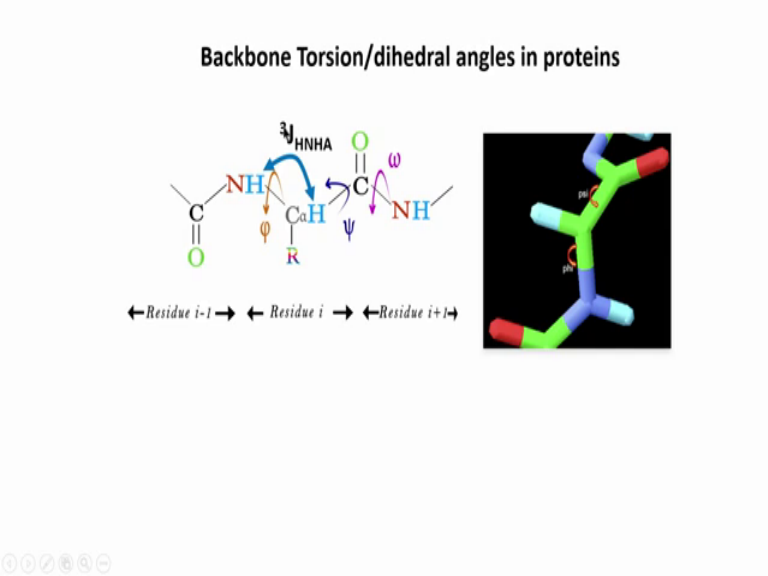
So, now, in recent years they have a lot of course, improvements what I have shown here this approach was a pretty old one. Now, in the recent times they have people have improved this method to even get more accurate prediction or accurate estimation I would not call it prediction; prediction is when you do not have NMR chemical shifts, but if you have NMR chemical shifts we can directly identify and that is published you can probably go to this website to get more detail idea of how CSI today is used for protein structures.

(Refer Slide Time: 05:40)



So, let us move on to the next structure secondary structure method that is based on the J-coupling between N and H alpha. This is HN H alpha, between HN there is a amide proton and H alpha proton. So, let us see what is that before that let us also look at what is what are torsion angles?

(Refer Slide Time: 06:00)



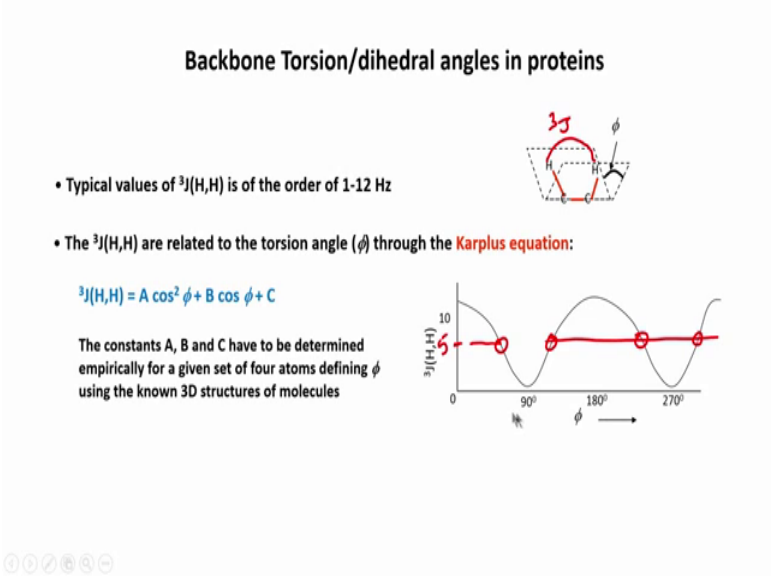
So, in a protein if you look at the backbone of a protein there are different atoms and it can define different torsion angles for different bonds. So, remember torsion angle is not same as bond angle so, bond angle is defined between 3 bonds and a torsion angle is

defined between 4 bonds means two planes. So, I would recommend you to strongly read this up in a textbook on any structural biology, any biophysics or NMR textbook and find out what is how do we actually define a torsion angle. So, many times people confuse a torsion angle with the bond angle which is completely not correct, completely wrong.

So, there now these are different torsion angles shown here. So, the very important torsion angle for secondary structure is these two; phi and psi. So, these two together tell help us to tell whether this amino acid. So, if you look at this is our amino acid I, these amino acid is alpha helix or beta sheet based on these values, but in NMR we normally measure only this value, we do not measure this value; we measure only this one. How do we measure this value we measure this value based on this coupling between this proton and this proton. So, if you notice this is a 3 bond separation between the two protons. So, we use this 3 bond separation we call it 3 bond J-coupling between HN and H alpha.

So, we will see now how we measure it and how we use it for finding the secondary structure of that amino acid. So, you can see this is more three-dimensional picture of how the torsion angle is defined. So, you can see this is a rotation around this bond. So, if these two groups here these play this a set of atom and these set of atom if they rotate around this bond then that corresponds to a phi rotation of this torsion angle similarly psi is shown here. So, one has to have a very firm grasp or from understanding of what is this torsion angle how do we define them? So, as I said we will use this 3 bond J-coupling with and we measure this J-coupling in proteins and then from there we will try to find out the value of this phi angle ok.

(Refer Slide Time: 08:25)



So, this is how as I said the phi is a torsion angle between 4 atoms. So, what do you do is you can imagine these 3 atoms as one plane you can see this dotted plane here and another three atoms another plane. So, when you take these two planes it looks like a book an open book and the angle between these open book it will define as a torsion angle this is some torsion angle, do not think of it as our protein torsion angle any general torsion angle is defined in this manner.

Now, typically in proteins and amino acids this value in this J HN H alpha value will be of the order of 1 to 12 hertz ok. So, this is a number which we should keep in mind because this is the range one should know for what is the value of a coupling for 3 bond this is called a vicinal coupling ok. We have seen in the first part of this course what is called a geminal coupling which is 2 bond and vicinal coupling.

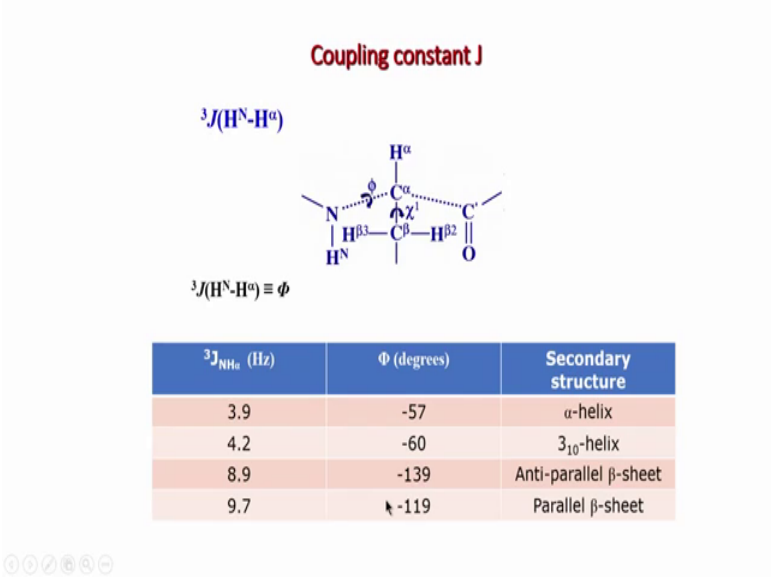
Now, this vicinal coupling constants that is 3 bond are very well known related very well known to the phi or the torsion angle for that. For example, let us say I measured this J-coupling between these 2 atom so, this is the 3 bond J-coupling. So, if I measure this 3 bond J-coupling between this hydrogen and this hydrogen that J value will be related to this phi value by this famous equation which is known as Karplus equation ok.

So, now if you look at if you plot this graph J versus phi it has kind of this nature and you can see that it is not straightforward because for a given J, let us say my J values somewhere here I can have multiple phi values which will have the same J values ok. So,

it is therefore, important for example, if I have a J value of 5 hertz it can actually be based on either it can be here, it can be this value phi this can be this or it can be this. So, there is not easy to find out the phi value straight away from J value, but still we get an idea that were the phi value could be, what kind of values we can expect to get.

So, that is how we use this method for based on J value we try to estimate what could be the phi value. So, this is kind of a approximation or vice versa; if I know the phi value I can estimate what is the J value, so, this is another option to use this approach.

(Refer Slide Time: 11:00)



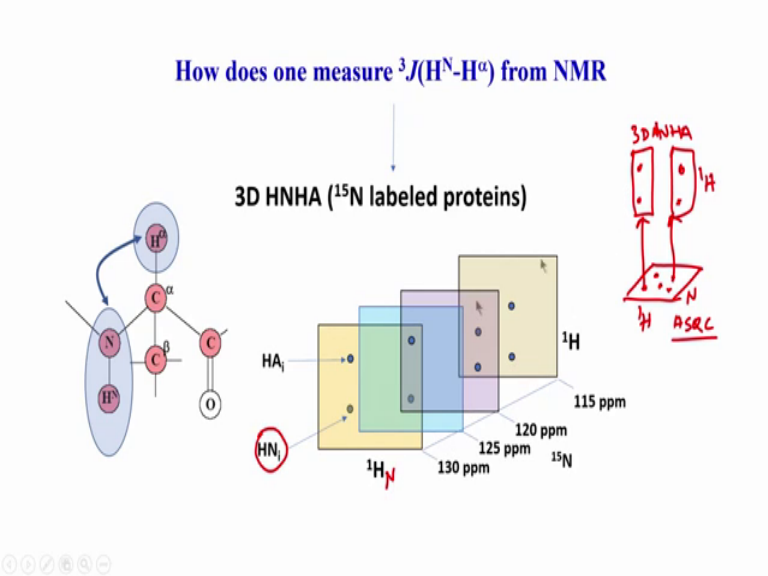
$^3J_{NH\alpha}$ (Hz)	Φ (degrees)	Secondary structure
3.9	-57	α-helix
4.2	-60	3_{10}-helix
8.9	-139	Anti-parallel β-sheet
9.7	-119	Parallel β-sheet

So, these are the set of values which tell us different secondary structure. So, we can see as I was mentioning parallel and beta anti parallel sheet there is a slight difference. So, if we look at this now let us start from the helix; if the J value is somewhere like 3.9 hertz 3 point this is an average of course, do not think that it will be exactly this value it will be somewhere between 3 to 4 hertz and that case our alpha the secondary structure could be alpha helix for this residue.

If the value turns out to be somewhere like 4 or 5 or 6 hertz sorry, 4 to 5 hertz then it will be 3 10 helix and somewhere between 5 to 7 or 8 will be random coil ok. The random coil is not shown in this table, but if you get a value of 6 hertz or 6 points say have somewhere between 5 to 7 hertz you can think of that particular amino acid is in the random coil confirmation, then very high value high higher to means towards the higher side 9 or 10 hertz tell us that they are in the beta sheet.

Now, there is a small distinction between alpha helix and beta sheet sorry anti parallel and parallel similar to distinction between alpha and 3 10 helix. So, you can see this typical average values again if the in the coupling constant is very high close to 10 hertz it will be somewhere can be thought of as parallel beta sheet and these are the typical phi values which are known in literature based on the three-dimensional structure of proteins. So, this is based on the high resolution x-ray structure people know that for an alpha helix what would be the average phi value and so on that is now shown here and it is correlated with the J-coupling values. This will also follow the Karplus equation which I showed in the previous slide.

(Refer Slide Time: 12:55)



So, the question now is how do we actually measure this J-coupling by NMR. So, J-coupling by NMR is actually done using an experiment known as 3D HNHA. So, this is typically done on the N 15 labeled proteins. So, what are we doing in this particular spectra? In this particular spectrum or in this particular experiment we correlate the nitrogen amide and this two atoms that is N 15 HN, that is a HSQC right, HN H alpha we correlate that to the H alpha of the same amino acid.

So, this is the point here we are correlating chemical shift or we are connecting the chemical shift of the HN then 15 and the H alpha of the same amino acid. This is very important point because we do not want to use go like HN CO CA where it correlated with the previous amino acid this is within the same amino acid.

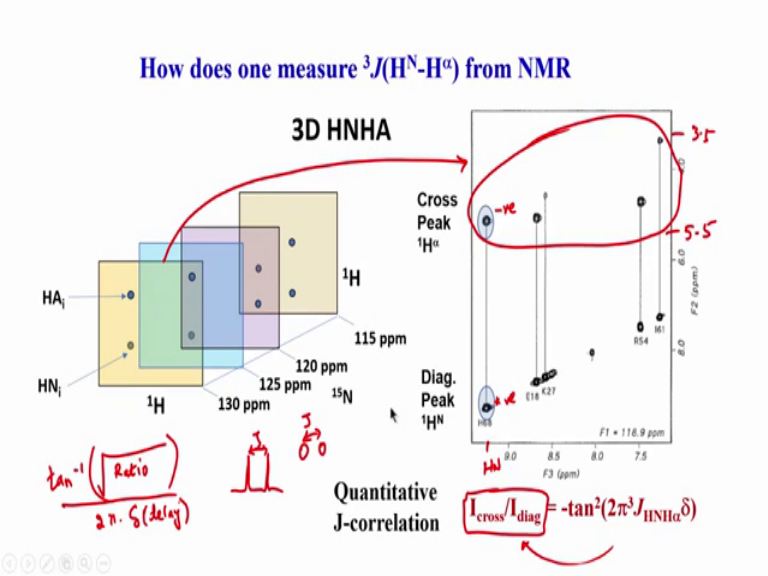
Now, the spectrum will come something like this; that means, in a three-dimensional manner we have seen this kind of a picture for other three -dimensional experiments. So, three-dimensional experiments consist of planes, in the third in one of their direction and other to gain in this dimension. So, this is a collection of planes along one dimension ok. So, collection of 2D planes is a 3D so, this is something we have been seeing repeatedly in the previous part of the course.

So, in this experiment HN H alpha for any given amide proton to nitrogen pair; so, this is like HS QC pair for any amino acid in the third dimension you will see two peaks; one is called the diagonal peak diagonal meaning ok, so, let me write the correlation. So, you can it is shown here so, this is basically this is called a diagonal peak, why because this and this are same this is also HN and this is also HN so, both are same. So, this is kind of a diagonal peak ok, but this is a cross peak and each amino acid like this will be separated in 3D plane by based on their N 15 value so, this is a 3D experiments so, it helps to reduce the overlap.

So, you can see here now for every plane for every amino acid I will for every amino acid I will get two peaks for a given NH and N 15 chemical shift and that is basically the HNHA. So, again if I want to make it more clear let us say that I have a 2D HSQC spectrum. This is HSQC 2D spectrum if I will get peaks like this in HSQC, now in HNHA if you go in this dimension and I will get two peaks here similarly for some other residue I will get two peaks in the 3D so, this is in the 3D HNHA ok 3D HNHA.

So, 3D HNHA we for every peak in the HSQC is so, this can be labeled as HSQC. So, for every peak in HS QC in the third dimension I am getting two peaks as shown here for a given a HNPN that is for a this is like one plane, this is another plane, this is third plane as many planes. So, in a given plane for a given residue I will get like this. Now, using this these two peaks will see now how we will calculate the HNHA for coupling constant.

(Refer Slide Time: 16:37)



So, this is what is shown here this is HNHA coupling constants, I am sorry the HNHA spectrum so, what is done is we do like this. So, this is for a one particular plane. So, let us say I take some plane I take this plane, but of course, in this plane I have shown here only one here, but it is possible that you may have some other amino acids also in the same plane ok.

Now, you go for each of this amide. So, this is each one of them is an amide residue amide chemical shift for each of the amide chemical shift you will see two peaks like shown here ok. So, one is the diagonal peak you can see it looks like a diagonal is going like this and another is a cross peak. Now, cross peak corresponds to the H alpha chemical shift value and the diagonal peak corresponds to the HN chemical shift value. So, you will see this will somewhere come around 7 to 8 ppm whereas, this H alpha this whole zone this zone is the H alpha chemical shift somewhere from 3.5 to 5 ppm ok. So, this is the alpha H alpha chemical shift values.

Now, what we do next is we actually integrate means we calculate the volumes of these two peaks. So, that is what is shown here by circle we integrate get the volume intensity of this peak, the cross peak and the diagonal peak and then we take the ratio. So, when you take the ratio of the cross peak to the diagonal peak then it is related by some formula to the J value and this is something which I will not be able to derive in this

course, but if you look at some advanced NMR textbooks you will be able to see that in that textbook.

So, what they is this comes from some equation from the pulse program as well from based on what is called as a transfer functions. Say it turns out that if you take the intensity or volume of the cross peak divide it with the diagonal for the same amino acid, ok. So, look here we are looking for the same amino acid. So, for the every amino acid we take the cross divided by the diagonal and from that value we will get that is corresponds to tan square into this or tan square of this, sorry.

Now, what I can do is I can invert this means from based on the intensity I can then take a square root of this and take tan inverse so, let me write down here. So, if I know this value so, that value I can take a square root of that ratio and call it as a ratio and then I take a tan inverse of this I am just inverting from here to here and then divided by 2 pi into delta; delta is some delay in the pulse sequence which is not mentioned here is a delay value in the pulse program which I did not show in this one, so, that delay value is delta.

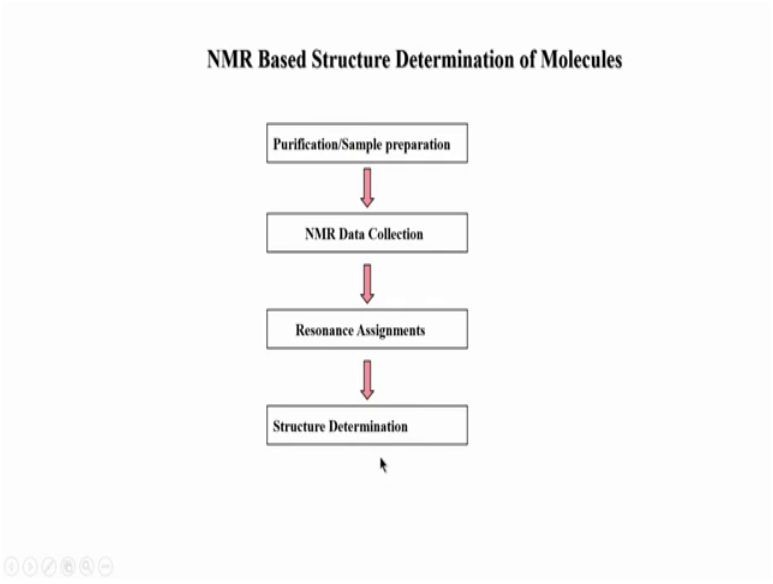
So, if I take this tan inverse and divide by this tan inverse sorry, I had to take a square root first then take a tan inverse and now here you see there is a negative sign. So, here your doubt will come why this is so? What happens is this intensity is actually negative this is negative intensity, this is positive. So, what happens is when you take this ratio; when you take this ratio the ratio will be negative because this is negative intensity, this is positive so, that this value will be negative. So, when this is negative and there is a negative sign here that will cancel and therefore, directly you will be able to take the square root because square root will be of a positive number only because the negative sign of here will cancel with the negative sign here.

Now, once you do that you can now take a inverse and take calculate the J value based on this formula that is you take the square root and then tan inverse. So, this is something which is routinely done and this kind of approaches in NMR in general. So, this is a very general terminology called quantitative J correlation. So, there are two ways to measure J correlation in NMR; one method is to look at the peak splitting. For example, if I have a doublet then I can use this difference and calculate J value, right.

So, this is called a simple spectral based J estimation, but here we are not doing that we are not taking the distance between this we are taking the ratio of the intensity ok. So, this is more this is like a quantitative approach this is not a this is also quantitative, but here we are using the separation between the doublet to measure the J coupling, but here we are using the intensity of the peaks to measure the J coupling. So, this is a very fundamental difference between the quantitative J correlation and other types of. So, this is how typically in 1D we measured like this, in a 2D we can measure with a contour ok.

So, this is how J values can be measured in a 2D, but in NMR remember there are therefore, there are two methods we either measure it from the spectra or we can measure using (Refer Time: 22:02). So, there are many experiments published in this area and similarly using this approach. So, you can also use J you can also measure HNHA J value using this approach, but this technique we have not discussed here. The one which I discussed today was more popular experiment which is based mainly on quantitation that is very much taking the ratio of the intensities.

(Refer Slide Time: 22:27)



So, this basically is brings us to the end of the secondary structure and as we move on we will see in the next class we will basically look at how now we can move from secondary structure to 3D structure. There will now start looking at NOE based methods and find out the tertiary structure of the protein.