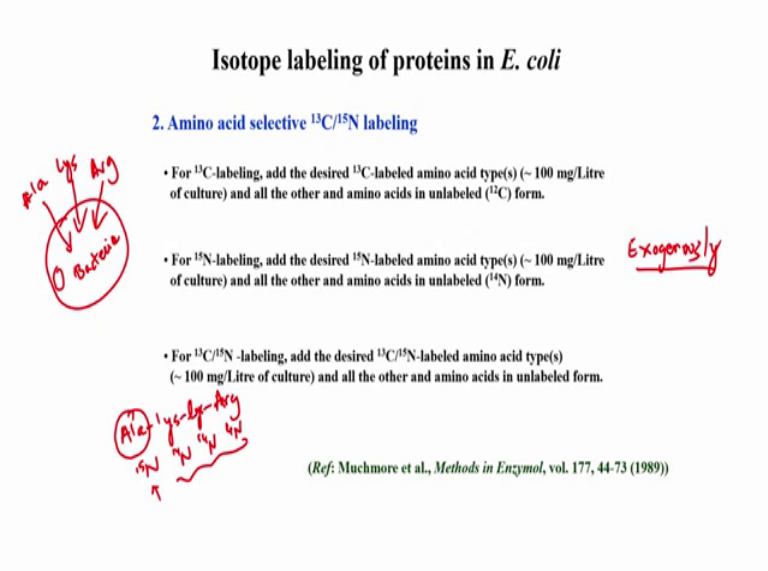
Multidimensional NMR Spectroscopy for Structural Studies of Biomolecules

Prof. Hanudatta S. Atreya
Department of Chemistry
Indian Institute of Science, Bangalore

Lecture - 25
Isotope labeling of proteins for NMR studies - Part II

In the last class we were looking at Isotope Labelling Schemes. We started looking at what is called as amino acid selective labelling, in which we want to selectively label a given amino acid type. For example, if you want to label A or B or C depending on what you have in your protein, and how is that done that was shown that it is done by adding the required amino acid. For example, let us say you are trying to label alanine, you add alanine in the required quantity which is about 100 mg per liter, that is in the C 13 or N 15 labelled form.

(Refer Slide Time: 01:00)



For all the other amino acids we add them in the unlabelled form. So, this is very important because we want only the desired amino acid that is alanine to be labelled and therefore, all the other amino acids are added in the unlabelled form. This is called exogenous because you are adding externally.

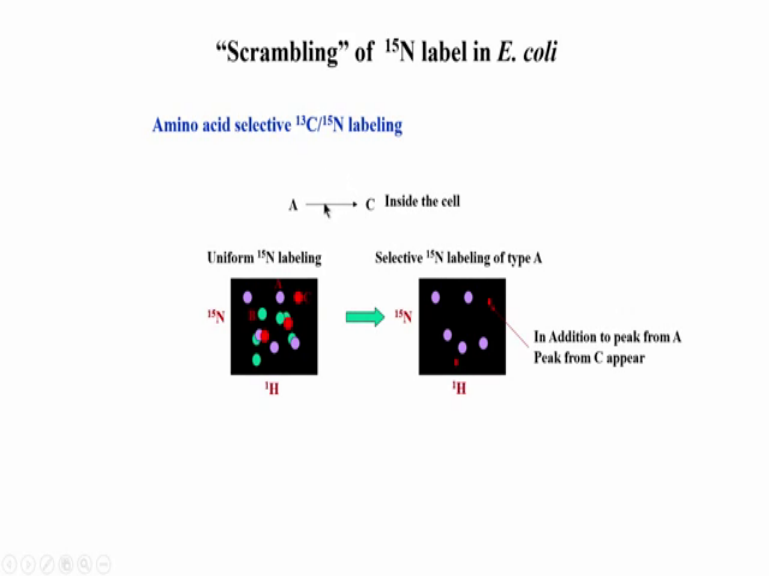
In case you want to do double labelling that means, both C 13 and N 15 labelling of an amino acid in that case we have to add the double labelled amino acid to the culture and up to the amount of 100 mg per liter of culture. And all other amino acids are now added

again in the unlabelled form. So, this is basically for a selective labelling of a particular amino acid type.

Now, let us say you want to say label two or more amino acid type then all those which you need to label should be added in the labelled form to the cell bacterial culture, and all the remaining which are not required to be labelled should be in the unlabelled form. Now, you see notice one thing here, we do not add any glucose are ammonium chloride which was done for uniform labelling, but for selective labelling no glucose are ammonium chloride is added and required.

Now, this is the standard scheme for a labelling, amino acid selective labelling. But there are some disadvantages of this approach, some drawbacks. Let us see what are those.

(Refer Slide Time: 02:24)



So, let us say in your cell inside the cell this amino acid A is getting converted to another amino acid C, ok. So, this can happen because there are lot of reactions happening in the cell and as I showed you a previous slide class, we looked at a biochemistry the biosynthetic pathway in which we saw that many amino acids actually derived from other amino acids.

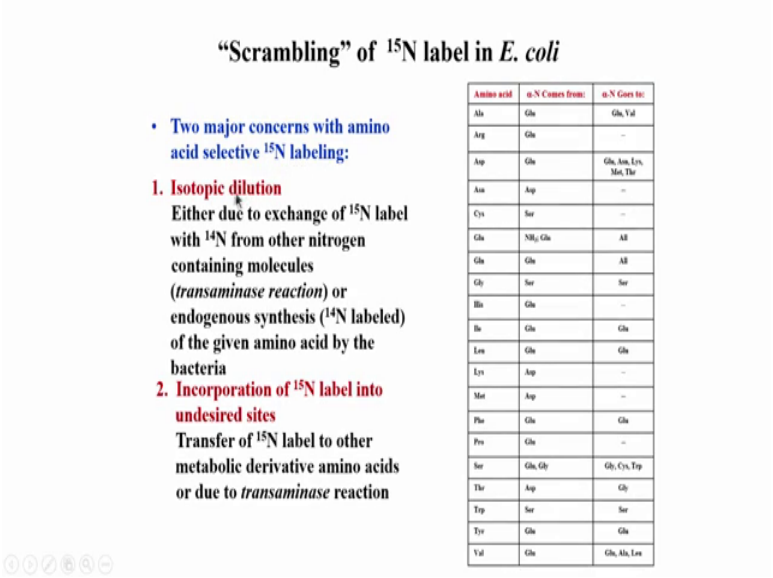
So, therefore, let us say that C an amino acid type C is coming from A in the cell and we get expect peaks from C and A in a regular sample. But here we want to label selectively N 15 for type A but because a is getting converted to C, the C peaks also appear in the

spectrum. We can see, so in addition to peak from a peak from C also appear in the HSQC spectrum. And why is this? Because A is getting converted to C. So, the labelling of A is carried over to the labelling of C. But the labelling is not very strong for C because we have added C from the external medium also in the unlabelled form. So, that is very high amount, we have added 100 mg per liter of culture and this conversion may not be very efficient definition. So, therefore, whatever C comes in the labelled form from A may not be as strong as the unlabelled form which is not seen here. That is the one which we have added from the in the medium.

So, therefore, the peak although C appears it is not maybe not be as strong as A, but sometimes it can be equally strong as A itself. So, you may mistake that peak actually corresponds to C, but in principle it belongs to actually C sorry you may mistake that it belongs to A but actually it belongs to the C.

Now, what is this whole thing called? This is called scrambling we use our isotope scrambling of N 15 label. Why label? Because N 15 is what we want to label the amino acid, but now this label is getting transferred from one amino acid to another amino acid because of some biochemical reaction inside the cell, and that results in this problem. So, this is very routine common problem in I selective labelling.

(Refer Slide Time: 04:43)



Amino acid	α-N Comes from:	α-N Goes to:
Ala	Glu	Glu, Val
Arg	Glu	–
Asp	Glu	Glu, Asn, Lys, Met, Thr
Asn	Asp	–
Cys	Ser	–
Gln	NH_3; Glu	All
Glu	Glu	All
Gly	Ser	Ser
His	Glu	–
Ile	Glu	Glu
Leu	Glu	Glu
Lys	Asp	–
Met	Asp	–
Phe	Glu	Glu
Pro	Glu	–
Ser	Glu, Gly	Gly, Cys, Trp
Thr	Asp	Gly
Trp	Ser	Ser
Tyr	Glu	Glu
Val	Glu	Glu, Ala, Leu

And I will show you now when chart where it is applicable to many amino acids. So, you can see for all the 20 amino acid listed here, they are lot of scrambling which can go

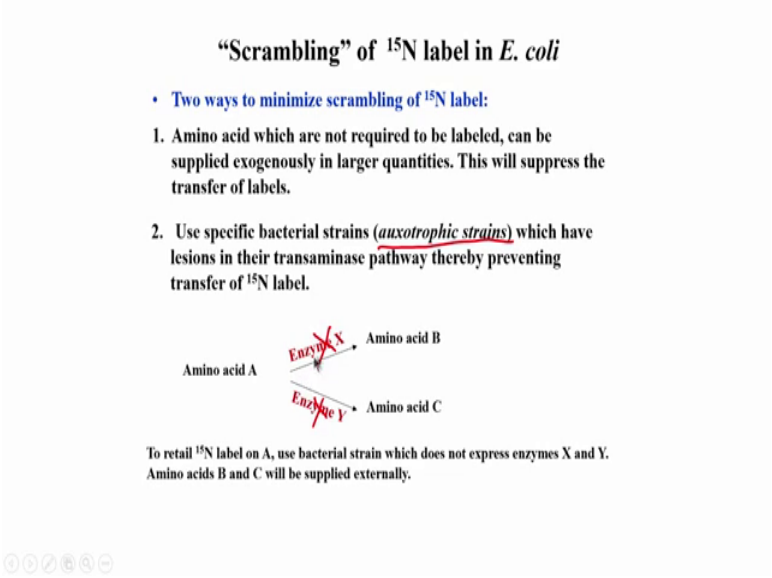
come from another amino acid. For example, alanine, the N 15 of alanine come from glutamine. So, if you try to label glutamine by N 15 your alanine will get also labelled.

Now, here you can see is another one from where it comes from and where it goes. So, you can see the N 15 of arginine is coming from glutamic acid. So, if glutamic acid is added in labelled form, arginine will also get labelled. Although we do not want that, we want only glutamic acid to be labelled but because of this scrambling problem it can also go to arginine. So, like this for all the amino acids we can see some of them are of course, very clean. For example, you can see here a cysteine does not go to any other amino acid, and it does not go to any. So, if you want to label cysteine then it is not difficult it will be clean. Similarly, asparagine does not go to any amino acid, which means if I want to label N 15 of asparagine and if I add N 15 label asparagine in the medium it will only an label asparagine it will not label aspartic acid or any other amino acid.

So, you can see this table is basically very standard table which is used. So, what is the major problems of N 15 labelling, that is selective labelling? One is isotope dilution because as I said N 15 label now gets into from another amino acid. And this is other thing is this transaminase reaction, so basically these two things table is listed here, ok. So, this reaction that is this which converts one amino acid N 15 transfers N 15 from one amino acid to another amino acid which we called as scrambling is catalysed by an enzyme. You need an enzyme in the cell to do that and that enzyme is called transaminase. So, this this transfer of N 15 label from one amino acid to another amino acid is very popularly by transaminase reaction.

So, question now is can we do something to stop this scrambling. Can we avoid this? Because we do not want extra peaks to come in the spectrum. The extra peaks which come in the spectrum will be reinterpreted by mistake as the peak which we intend to label, but in reality that is not the intended label it would have come from the scrambled amino acid and therefore, we may make a mistake in assignment or identification. So, therefore, they should be something done to avoid this scrambling and that is what we can see next. How we do that?

(Refer Slide Time: 07:33)



So, there are two ways to minimise scrambling. Of course, we cannot completely avoid scrambling because that is not possible of course, you can minimize. So, that the extra peak which comes in the spectrum will be a very small intensity. So, how do you do that? So, amino acid which are not required to be labelled can be supplied exogenously larger quantities.

So, remember I said in the previous few slides that we apply supplied 100 mg per ml mg, 100 mg per liter of the labelled amino acid per liter of culture, but in this in the remaining amino acids are also added to that level 100 mg per liter. But we can increase the amount we add of the unlabelled amino acids which we do not want to be labelled. So, if we add a larger amount to the cell and the culture the bacteria will have larger quantity taken up into the cell, and this will help in suppressing or reducing the transfer of the N 15 because the N 14 which is coming from the unlabelled will be retained in a larger population than the N 15 which is coming from inside the cell because of scrambling.

So, basically this is one way simply increase the unlabelled amino acids in the medium which you do not want to be labelled. Or the second option is you can use special bacterial strains. Now, this is something which you will have to buy commercially or borrow from biology labs where they have what is called auxotrophic strains. Now, this strains they have lesions, lesions means they the transaminase reaction does not take

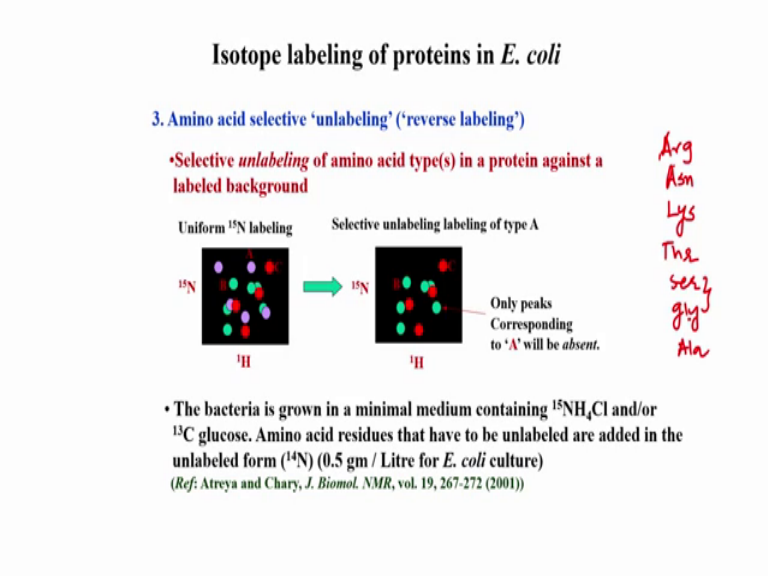
place as you as the required and therefore, that helps in preventing transfer of label. That is it helps in preventing a conversion of one amino acid to another amino acid. So, this is what is listed here.

Suppose I have an amino acid A which I want to label N 15 label, but now because of this transaminase two enzymes it is getting converted to B and C they are not the amino acid the N 15 is getting transferred because of this transaminase. So, remember the word transaminase, means amine something to do with the amino group which is getting transferred. So, the NH 2, NH 3 from one amino acid gets transferred to B and therefore, that nitrogen is also transferred labelled.

Similarly, there may be another enzyme which goes to C. So, now, suppose I want to selectively label only A I should stop these two reaction, and by doing this I if I use an auxotrophic strain in which these two enzymes are not present in good amount because of some mutations, then I can actually stop it from working ok. So, I can I can have a bacterial particular strain which will basically stop this reaction because of this particular strain, and in that strain I can now use that and if I grow my cells I will be able to selectively label A without having the problem of getting contamination or scrambling to B and C.

Now, the question come may come to your mind then how should B and C be made in the cell. It need not be actually the N 15 for these need not be made from here, because we are supplying the N 15 N 14 amino acid B and C in the medium. So, when you are supplying those amino acids the bacteria should be able to take up these amino acids from the medium and it may not need to make from A. So, therefore, these enzymes are actually not needed. So, auxotrophic strains are those strains which are specifically tailored or engineer to achieve this objective. So, that can be also used.

(Refer Slide Time: 11:26)



Now, there is an another approach to selective labelling which is actually not a labelling approach it is a unlabelling approach. So, this is something which is being done in recent years. So, this is actually reverse opposite of labelling selective labelling. So, let us see what happens here.

So, in this case in this labelling scheme see this is the standard HSQC. So, this is an example I showed you earlier, then let us say you have 3 amino acid types A, B and C in your protein and each of this A type is giving peaks, each of the B is giving N, C are giving the peaks.

Now, what we did in the selective labelling? We selectively labelling approach we added A in the labelled form to bacteria and B and C were added in unlabelled form. So, because of that B and C were not seen in the spectrum and A was there in the spectrum only. But here we do the opposite. We supply, so we do not actually supply any amino acid we try to achieve that let us say that I want to suppress A, I want to unlabel A, but B and C I want it to be labelled.

So, how do I do that? I can supply A now in unlabelled form remember this is exactly opposite to what we did in selective labelling. There we supplied A in the labelled form here we are going to supply A in the unlabelled form, and because of that the A will be absent, but B and C are going to be in the labelled form. But are B and C going to be supplied in a labelled form that we will see. But the idea objective of this approach is to

get rid of one of the amino acid type, one or more and keep the and see the remaining ones. So, you can see here only peaks corresponding to A will be absent because I have selectively unlabelled A from the spectrum.

So, how do we achieve this? So, this is the protocol shown here I can refer to this paper which gives more details. So, the idea here is that a bacteria is grown in a minimal medium. So, this is the standard minimal medium which contains. Now, ammonium chloride N 15, or C 13 glucose and or C 13 glucose.

Now, what we do is additionally, this is the standards, so up to this line the first line if you read this is the standard protocol for uniform labelling. But for unlabelling now what we do is we add an amino acid residue that has to be unlabelled. For example, here A want we want to unlabel A, we add A in the unlabelled form to the medium. And how much do we add? We add about 0.5 grams for liter of culture, ok. And what happens with other amino acids? The other amino acids we do not have to do anything. Why? Because it comes from already this label here and this label here. So, bacteria takes N 15 C 13 into the cell and then starts making all these A B C D amino acids, but A it does not make because A is already coming from outside. So, when you are supplying A from outside it does not have to make A in the cell, it only makes B and C. So, the synthesis, biosynthesis of A is cleanly removed suppressed by adding A in the labelled unlabelled form in the medium, ok.

So, because you are adding a large quantity that quantity is huge compared to what we did in labelling selective labelling. So, therefore, the bacteria very is able to suppress the 9 14 which is coming from the now; it is able to basically take up that from the cell from outside and it does not make it inside the cell.

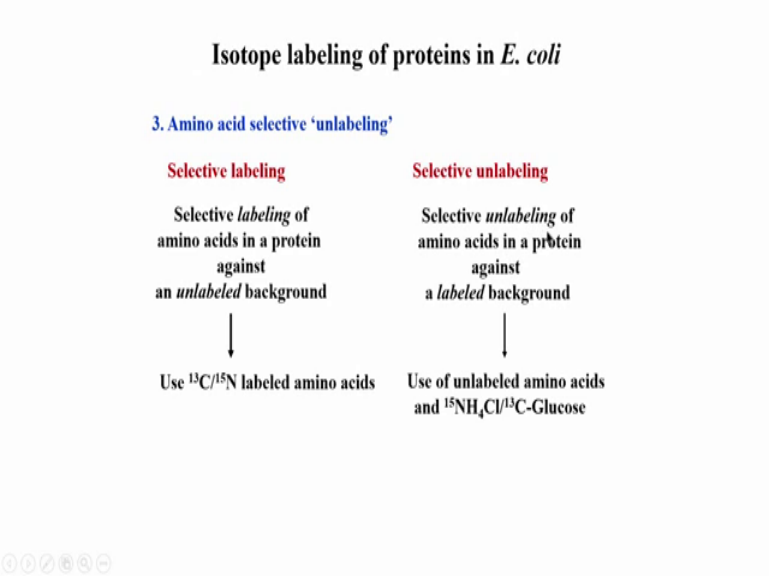
Now, the question you may come to your mind is scrambling not a problem in this case, because I would expect suppose B and C are converted to A, then the whatever B and C labels are there will also end of label on the A. And then we will end up seeing A in the spectrum. Even though we have supplied large quantities of N 14 A in this medium it may still happen that B and C are converted to A by the transaminase reaction and because of that it can happen that A gets now labelled and starts appearing in the spectrum. Is that possible or not? Yes, it is definitely possible. The scrambling problem is the same as in selective labelling. So, this is similar to what we saw in this case.

So, for example, if you look at this chart here, if I try to unlabelled aspartic acid for example, if I want to unlabel aspartic acid, I will add aspartic acid in the unlabelled form into the cell like we did, but then that will also get converted to asparagine inside the cell. So, even though the cell is trying to make this labelled form of asparagine in the cell in because of I am adding unlabelled aspartic acid outside it is getting converted to this because of this transaminase. So, my asparagine also will start getting unlabelled.

So, selective unlabelling is also is similar problem has similar problems like selective labelling. But a few amino acids can be very easily unlabelled without any problem. So, let me list those amino acids which can be labelled without a problem. So, one is asparagine, lysine, threonine, serine, glycine, and alanine. Of course, one thing is these 3 are inter, these two are inter convertible. So, if I unlabelled serine, glycine also gets converted to serine gets converted to glycine and vice versa, and threonine, lysine, arginine these are very conveniently you can unlabelled without any problem. We can also unlabel arginine.

So, these amino acids are one which do not scramble. So, if you want to do selective labelling or selective unlabelling for both the the schemes these amino acids are very preferable because they try to, they do not interconvert to other amino acid of course, serine goes to glycine. So, these two are inter convertible but threonine or lysine, asparagine, arginine, alanine, to some extent glutamine, glutamine actually is not a good idea because glutamine unlabelles many other amino acids. Similarly, in selective labelling also we do not use glutamine for selective labelling because it scrambles very rapidly to other amino acids. So, these are the few amino acids which can be safely labelled or unlabelled selectively.

(Refer Slide Time: 18:10)



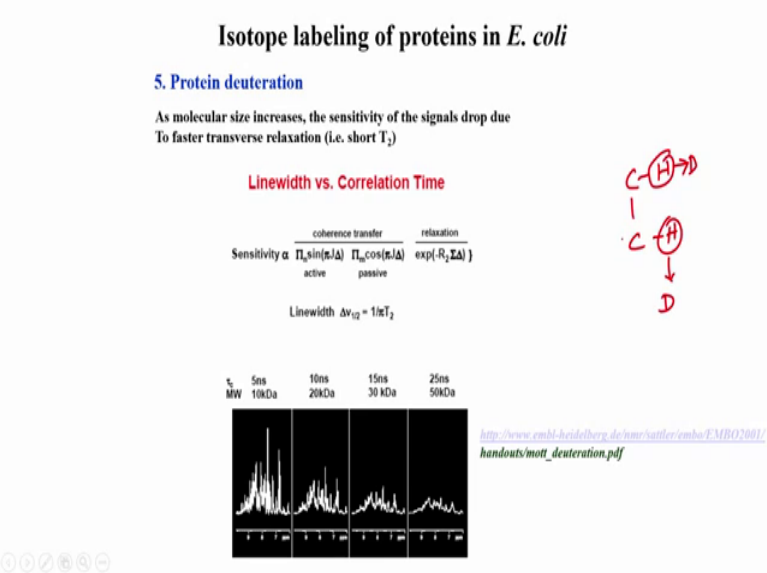
So, now let us see the advantages and disadvantages of these two approaches, that is unlabelling and labelling. So, the advantage or differences, let us look at the difference. The difference here is a selective labelling of an amino acid takes place in the protein, against means the remaining amino acids are unlabelled. Whereas, here a particular amino acid type is unlabelled selectively but the remaining amino acids protein are labelled. So, it is a opposite or reverse of this approach.

Now, here we have to use C 13 N 15 labeled amino acid because as I said this is supplied exogenously or externally in the medium and therefore, if you have a C 13 if you want to do N 15 labelling or C 13 labelling those particular amino acids you will have to buy commercially and put it in the medium. Here we do not do that. Here is we do not have to add any labelled amino acid. Of course, we have to add unlabelled amino acid. So, we use unlabelled amino acid in the medium, but we are the remaining amino acids are not added they all come from these two components which are added in the medium.

So, now if you look at the cost wise, the cost of labeled amino acid is pretty high if you want to commercially by this. Whereas, cost of these two components are not very high as compared to the relatively, relatively these are much less expensive. So, overall this of course, unlabelled amino acids are I am not even counting them because they are hardly any price is very cheap. So, these are almost can be ignored. So, only cost involved is this part here N 15 this or N 15 amino acids. So, this approach overall therefore, becomes

much less expensive compared to this approach or you can, so therefore, depending on the requirement and the cost and so on one can choose which approach one has to take for labelling.

(Refer Slide Time: 20:07)



So, let us move on to the next labelling scheme that is deuteration, protein deuteration. So, deuteration what is this let us see now. Before we go into the details first you see this picture here which is been taken from this website. One can go to this is the deuteration is a pdf.

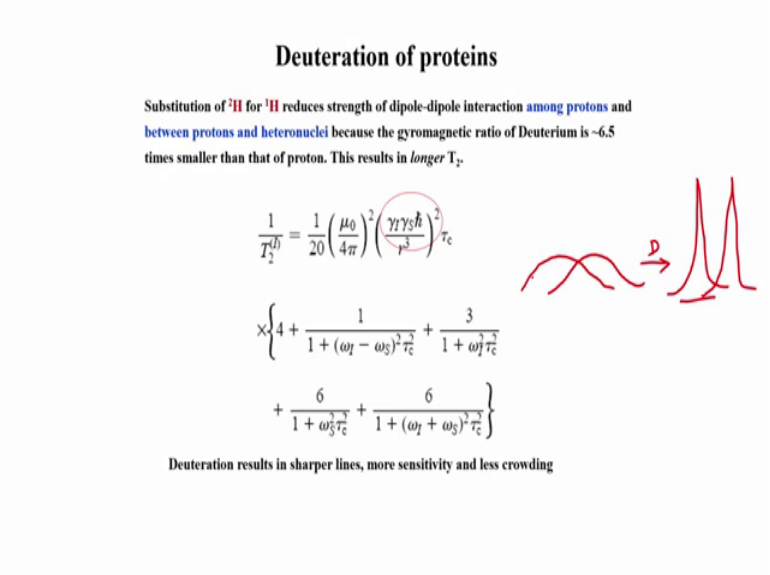
What it shows here is that as we increase the molecular weight from 10 kD to 20, the signal to noise or sensitivity is going down which is why is expected because T 2 relaxation of the protons they become shorter and shorter, and when T 2 is short the lines become broad and when the lines become broad the sensitivity goes down, ok. So, this is very simple idea there as increase the molecular weight the size is going down. Now, if you notice this tau C here this is simply the using that formula which I showed you earlier the tau C nanosecond is molecular weight in kilodaltons divide by 2. So, 50 kD will be 25 nanoseconds.

So, now, with this because of this now, the question is how do we solve this problem; if the sensitivity is going down so dramatically can I do something to improve the sensitivity of this or resolution. Because remember I have mentioned it is a broadening effect the lines are getting broadened because of the shorter T 2 of the proton. So, this is

what is mathematically shown here, that you have basically relaxation which dominates, relaxation start dominating and these are simple coupling constants. So, this may not change with the size of the protein. What changes with the size of the protein is this part here relaxation and that is because of this R 2 that is T 2 and that can depends on 1 over T 2, R 2 is 1 over T 2. So, therefore, to improve the sensitivity the method of deuteration is done.

What are we do in deuteration? In deuteration every nonexchangeable. So, let me draw here every nonexchangeable protons. So, for example, let us say you have this in your amino acid we replace this with deuterium, ok.

(Refer Slide Time: 22:29)



So, because of this deuterium what happens is that is what I will show you in the next slide, that because of the deuteration effect this the d 2 relaxation changes dramatically. So, you see this equation here which we are not going to derive, this is given in many papers and textbook but this basically this equation is telling you that the T 2 of a proton, it or carbon and proton any T 2 of any molecule any atom or nucleus depends on the dipole this is coming from the dipole interaction of two nuclei.

So, for example, let us say I am looking at carbon. So, the T 2 of carbon, so let us say I is carbon T 2 of carbon depends on the dipole interaction with proton or another carbon. So, therefore, if I remove one of them with deuteration like here substitution of 2 H, that is deuteration for a proton will be bring this term very down rapidly. Why is that?

Because now the gamma of deuterium is 6.5 times less or smaller than that of proton. So, because of that when it is less this factor goes down but remember there is a square here. So, it goes down by 1.65, 6.5 times square is even higher. So, it goes down dramatically this term. So, when this goes down dramatically T 2 will go up because T 2 is in the denominator here.

So, by the simple mathematical concept you can see that if I reduce this gamma value. So, suppose I am looking at I, so I should reduce the gamma of S if I reduce the gamma of S by replacing suppose this is proton I change it to deuterium I reduce it by almost square of this term. Ideally of course, we will not achieve this much but yes it is possible to achieve very high reduction in this term by replacing a deuterium and that in turn increases T 2. So, T 2 becomes long. So, when T 2 becomes long the lines become sharper, lines meaning the peaks and therefore, it results in more sensitivity and less crowding. Crowding because overlap is reduced and because overlap is reduced I can get more.

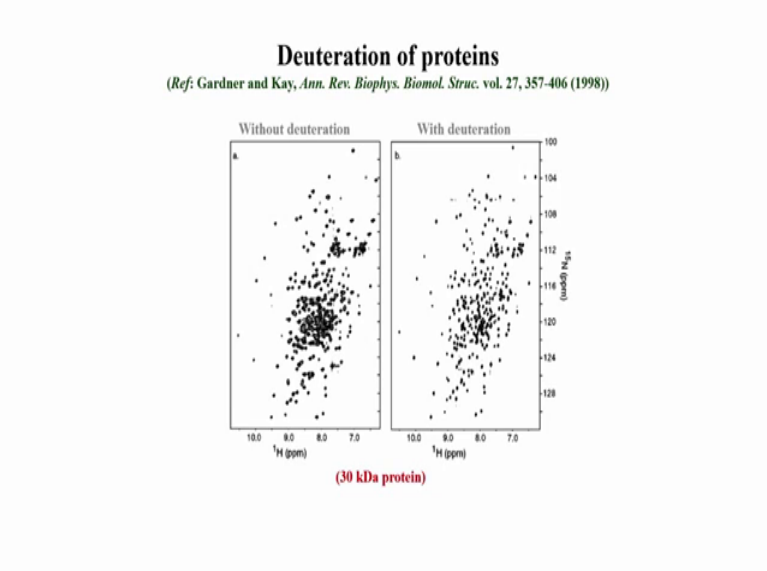
So, we can think of it like this, suppose the line earlier was broad line because of deuteration it becomes a very sharp and it becomes stronger. So, you can see this is the effect here. So, it reduces. So, if I have another peak here then it will be very sharp true lines. So, you can see I can resolve the peaks, will be better compared to here not only that I will get much better signal to noise compared to when it is broad.

So, deuteration that is replacement of deuterium or replacement of proton by deuterium again remember we are looking at nonexchangeable. So, we are not considering NH oh protons which when you put it in water they exchange rapidly with solvent. So, therefore, we are trying to look at I mean we are talking about deuteration of the protons that is CH protons. So, here again I go back to this slide we are looking at deuteration of this here, ok.

Now, depending on this there are a lot of schemes available. For example, sometimes you may not want to completely deuterate all the hydrogens, you may want to partially means you want to keep this hydrogen as deuterium or replace it with deuterium but you may not want to touch this hydrogen for whatever reason we will see that. So, these are called different labelling scheme that is labelling partial deuteration, fractional deuteration, random deuteration, uniform deuteration and so on. So, there are different

types of deuterations which we can do which we will see shortly and that will depending on your application and that will help in resolving the peaks and reducing the overlap.

(Refer Slide Time: 26:29)



Now, let us see the effect of this on N 15 HSQC for a protein. This is been taken from this paper this is a review article this is very old one. But you can see there how dramatic is the effect. So, this is the 30 kilodalton protein. So, you can see without deuteration the HSQC now N 15 proton. So, by this time you will be now familiar with these kind of spectra, that what kind of data we get, what kind of information we get. So, this is an N 15 HSQC of this protein, and you can see that here it is basically the peaks are very broad and crowded, but the moment you deuterate the signal to noise is increased because the peaks are gone up in intensity and not only that the peaks are very sharp and you can see they are much better resolved compared to what you see here.

So, deuteration therefore, helps in resolution, sensitivity and the and therefore, preferred for a large proteins. Typically, we do this if you are in 25, 30 or 50 kD or 70 kD protein when you go to higher level that is about 25 kD proteins, 25 kilodaltons which is roughly 200 220 amino acids, if we have that kind of protein. Or if you may have a multimer you may have a smaller protein you may have a 12, 15 kilodalton protein let us say but it may be forming a dimer or it may be forming a timer. So, what happens is a 15 kilodalton protein when it forms a timer it is actually a 45 kilodalton protein because there are 3 units or 3 chains or 3 protein molecules together. So, in such cases also many deuteration

to reduce the line width, because even though the number of peaks are less but still the lines will be broad sensitivity will be low. So, you can deuterate the protons this protein and you will be able to get much better line width.

So, in the next class, now we will move on to see different deuteration schemes. As I mentioned the deuteration itself there are several variations. We will see how different variations are there, and how does it help to further you get the better information.