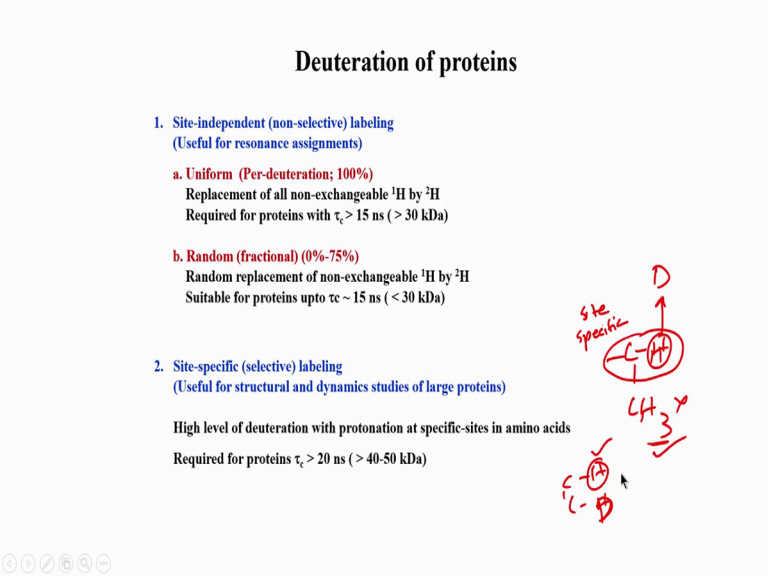
Multidimensional NMR Spectroscopy for Structural Studies of Biomolecules

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

Lecture – 26
Isotope labelling of proteins for NMR studies-Part III

In the last class we looked at how deuteration helps to improve the sensitivity and resolution and we saw this example of this 30 kilodalton protein where, the HSQC we gets much resolved better resolved then when it is not deuterated. So, now, let us look at different schemes which we need for deuteration which is shown here. So, for example, you can do like we did for selective N 15 labelling you can have uniform deuteration. So, when you say uniform deuteration we use the word per deuteration. So, what do we do in uniform deuteration, we uniform deuteration we replace all the non-exchangeable hydrogen's with 2 H.

(Refer Slide Time: 01:05)



So, that is the key here, the key here is all of them not just one a few of them all the non-exchangeable; remember this is a keyword here are replaced with deuteration a deuteron deuterium. Now, when is this required? As I mentioned this is required when the protein is 25-30 onwards we do this. So, this is 100 percent deuteration, but sometimes you may not want to do 100 percent deuteration because, your protein may not be that big number

1. Number 2 is that when you replace proton with deuterium remember we lose this protons there are no protons in the protein now.

So, if there are protons in the protein the experiments like NOESY will have a problem. Because remember NOESY in I showed you earlier we saw in the 2 D part. The 2 D NOESY and we will see 3 D NOESY as well; all they rely on is hydrogen to hydrogen proximity distance. But, if there are no hydrogens in the molecule then how do I get the distance information at all. So therefore, we if you completely replace a proton with deuterium you cannot have the structural information, the proximity information or the distance information which comes from the NOESY experiment.

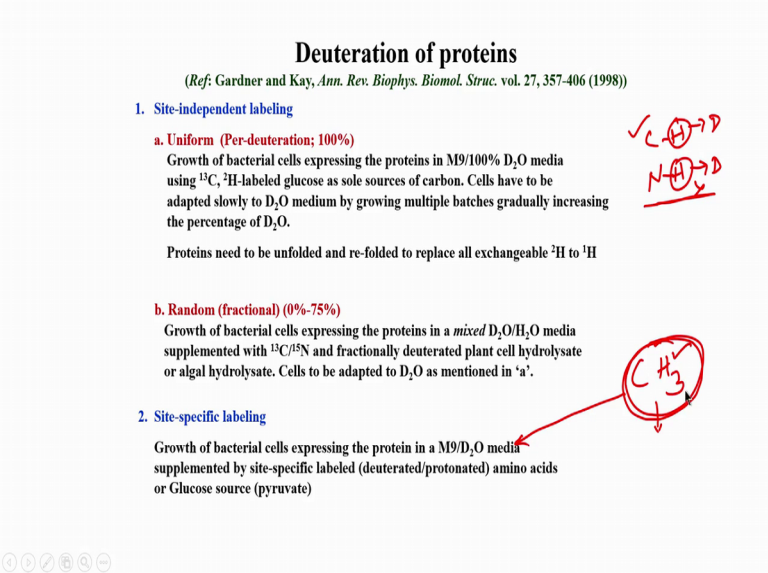
So, what is done sometimes is to have a compromise and what is the compromise; you do not directly label you do not label 100 percent with deuterium or you do not replace 100 percent with deuterium you label roughly 50 percent or 75 percent. And, that is sufficient to get enough hydrogens back because this is now randomly done, randomly means is not like systematically removing only a certain type of hydrogen we are replacing all hydrogen randomly with deuterium. So, there will be some hydrogen's which are left without replacement and they will be useful for you NOESY. Because, in NOESY remember again we read critically the hydrogens ok. So, this is typically done if the protein is not very big, but not small as well; that means, is of somewhere around 20 kilodaltons.

It can also be done for larger proteins, if you want to solve the structure with NOESY. Now, there is another option we can do with deuteration is called site specific which means a particular type of CH that is non-exchangeable proton CH there we choose for replacement with deuterium or a particular type of CH is left as CH and remaining we change to deuterium. So, this is something which I can illustrate here suppose I have two, 1 type 1 CH 1 CH 3; we will take CH 3 case in the next slide. There suppose I want to replace this proton with deuterium ok, but I do not want this to be deuterated, I want to keep this as proton itself ok.

So, this case called site specific site specific deuteration. Why? Because, this particular site is what I want to deuterate and I am not going to I do not want to deuterate CH 3 or it may be the vice versa. I have another system where I have CH CH CH 2 CH I may want to make this as remained as proton, but I may want to place change this to D. So,

this kind of sites specific deuterations are sometimes needed for very large proteins like 40 to 50 kilodaltons and we will see now what are the different type of site selective deuteration that is done in proteins.

(Refer Slide Time: 04:39)



So, one of the thing most important is how do we actually do the deuteration. So, this is shown in this slide. So, before we go into details of site specific deuteration; let us see in general how do we deuterate a protein, what protocol is used. So, first thing is you start the growth of bacterial cells like we saw in the uniform labelling system is similar to that, you grow the cells in a very specific medium that is M 9 minimal medium. The same medium which I said earlier, but the difference here is that media is now prepared in D 2 O 100 percent D 2 O; you are not preparing the media in H 2 O.

H 2 O is what we normally use in the uniform labelling, but in deuteration we need to use D 2 O medium that is one thing. Secondly, the glucose which we now use also has to be deuterated glucose. So, it cannot be protonated because what and why are we doing all these because you want every hydrogen to be replaced with 2 H; that means, I should not have any hydrogen in the protein. The protein should be entirely coming with C 2 2 H that is deuteration so; obviously, wherever I should I have hydrogens I should change that to deuterium. So, this is a carbon 12 C 13 proton labelled normally protonated, but now we use deuterated glucose.

Now, the one problem with this is suppose I directly start growing my cells in the 100 percent D 2 O media with this glucose it will not grow so well. Why? Because, bacteria needs to adapt or need to evolve to grow in the D 2 O medium. So, it has to be slowly acclimatized, slowly it has to be made adapted to the medium. So, what we do normally is we do not directly take 100 percent D 2 O we take slightly increasing percentage of D 2 O. So, initially we grow the bacteria in 30 percent glucose then make it 50 percent 80 percent then finally, do it in 100 percent.

So, the initial bacterial cultures are not directly grown in the 100 percent because, it may not be used to that medium. Remember D 2 O is not simple it say for example, human beings cannot drink or grow in this medium, it is very poisonous its dangerous. But, bacteria can somehow survive they can adapted very well compared to humans. So therefore, they can grow in D 2 O not so rapidly, but at same time it can still manage. So, how do we help it to manage? We slowly increase the percentage of D 2 O so, that it goes up and then finally, when it is 100 percent D 2 O it can grow and then the entire bacteria now takes up this from the medium.

So, there is no hydrogen in the medium. So, if you see here whatever carbon glucose we are using has 2 H. So, there is no hydrogen here proton here and there is no hydrogen or proton here. So, everything is deuterated now; that means, there every hydrogen in the protein is deuterated not just non-exchangeable hydrogens, but also the exchangeable hydrogen. That means, both CH as well as NH both of them are replaced with are basically deuteraed, but remember our interest is only in this. We do not want this to happen ok, we do not want a non-exchangeable let a exchangeable hydrogen we do not wanted to be deuterated.

But we cannot avoid this happens in the cell. So, how do I get back this hydrogen? I have to now put it in a what I can do is I can unfold the protein because, if I unfold the protein the entire hydrogen is exposed. So, an entire hydrogen is expose the protein now will start coming back to deuterium sorry hydrogen. So, this is what we do we have to unfold or denature the protein and when we denature the protein the protein opens up. And, when the protein opens up all the hydrogen bonded exchangeable amide which are now deuterated will get slowly exchange to back to original hydrogen.

So, we do these unfolding and refolding things inside in a water sample. So, we take the protein, extract the protein, we purify the protein after purification we follow this step of unfolding and folding. So, this unfolding folding is mainly for these atoms for this hydrogen's we want them back. So, we want the deuteration to go away. So, whatever is deuterated now will go back to proton because, we are dissolved this protein in a water sample H 2 O, but the growth and everything is D 2 O. So, we have to keep with this in mind; the initial growth, extraction, purification is in D 2 O but finally, we are exchanging it back to protons.

But this remains deuterated this will not get exchanged these are called non-exchangeable hydrogens. So, the non-exchangeable hydrogens do not exchange and therefore, whatever you do whether you do the unfolding, refolding this will not change whatever deuteration has come to this side will remain deuterated. So, this will remain deuterated ok, but this will get exchanged to hydrogen. So, this is typically what is done to get the exchangeable hydrogens and that to protons. Now, in random fractional deuteration how do we make the protein. Here we do the same thing we grow the bacteria in a now a mixture of D 2 O H 2 O.

So, you see this is no longer a D 2 O pure D 2 O medium like here it is a mixture means, you can decide how much you want to label. Suppose you want to do 50 percent labelling, that is random 50 percent deuteration what do you mean by 50 percent means; 50 percent of this hydrogens randomly are deuterated. So, if you want that kind of a deuteration you grow this bacterial culture let us say little higher percentage, not in 50 percent you can grow it in 70-75 percent and the remaining 25 percent will be water and 75 percent will be D 2 O.

So, is a mixture now, once you mix and then you have to give the supply the cells with these source of carbon and nitrogen. What is the source? Remember the sources glucose and 3 13 glucose or N 15 ammonium chloride. So, if you supply these two glucose and ammonium chloride, but the media is a mixture then you will end up with a protein which is basically partially randomly deuterated. So, this again here the cells have to be adapted for example, if we want to grow in 75 percent D 2 O you can directly put you cannot directly put the bacteria in that medium because, it will start dying. So, it is not able to adapt. So, we have to slowly increase the percentage.

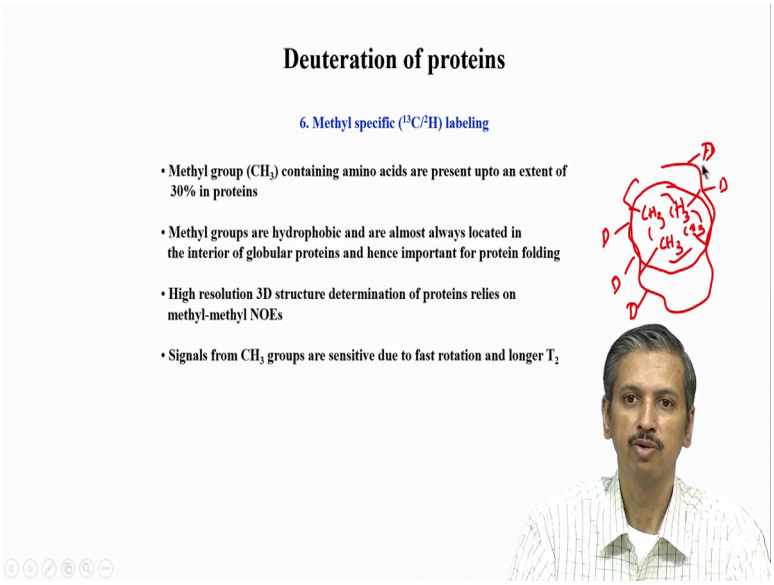
So, start from 5 percent 10 percent 30 percent and so on. So, this has to be optimised for every protein it may depend on what kind of protein you are expressing, what is the yield so on. But, this media adaptation is very important step when we are deuterating a protein. Now, one thing is D 2 O is expensive solvent, it is not as cheap as water. Water is available in NOESY very abundant in nature, but D 2 O is not. So, when we grow the cells in D 2 O typically what we do is the cells which after the cells are grown the media which we get back is actually recycles. So, we recycle D 2 O because, D 2 O as such if you keep buying is not so, easy it is not it is not cheap it is expensive.

So therefore, we take the media which is remaining after the cells have grown and so, that is called as a spent medium. The spent medium is now recycled by several distillation steps back to the D 2 O; now the D 2 O comes back and again you can make a new sample. So, this is the recycling approach, but again that the moment you do some kind of a recycling your D 2 O gets contaminated with H 2 O; that means, it may not be 100 percent H 2 O. So, you have to do a series of distillation and get it as pure as possible. Now site specific labelling how do we do that in bacterial culture.

This is again very important one know. So, as I said site specific meaning you want a particular site CH 3 for example, or aromatics to be deuterated or protonated. So, if we want to do that then you have to grow the bacteria again in this type of a medium. But, now what you have doing is the in medium we are supplementing it site specific labeled amino acids. When you do that in the site specific labelled amino acids the glucose or the pyruvate source is actually not required. Or you can use a particular glucose source like pyruvate or you can use a particular labelled amino acid and that now becomes incorporated in the medium.

So, let us say let us for example, let us let me suppose you want to keep the methyl group protonated you do not want to deuterate. So, if you do not want to deuterate methyl groups you have to supply the methyl amino acids in the medium. But, the medium contains D 2 O. When the medium contains D 2 O the remaining hydrogens will all be gone they will be now deuterated, but the methyl's will remain protonated because, we have supplied that methyl group in a protonated form.

(Refer Slide Time: 14:21)



So, this is what is shown why do we actually do methyl protonation. The methyl protonation is very useful because, methyl groups are actually very important in a protein. So, if you look at a structure of a protein; let me show you the protein structure like this. So, suppose there is a hydrogen here and there is a hydrogen normally the CH 3s are located in the core of the protein. So, this is for a folded protein. So, we can see this proton to proton they are very close by. So therefore, they have very good distance information for my structure determination. So, if I want to determine the structure of this protein and this distance between these different methyl's are very useful for me for structure determination. Therefore, it is better to keep this as protonated not to deuterate this because, if you deuterate this you will lose this information.

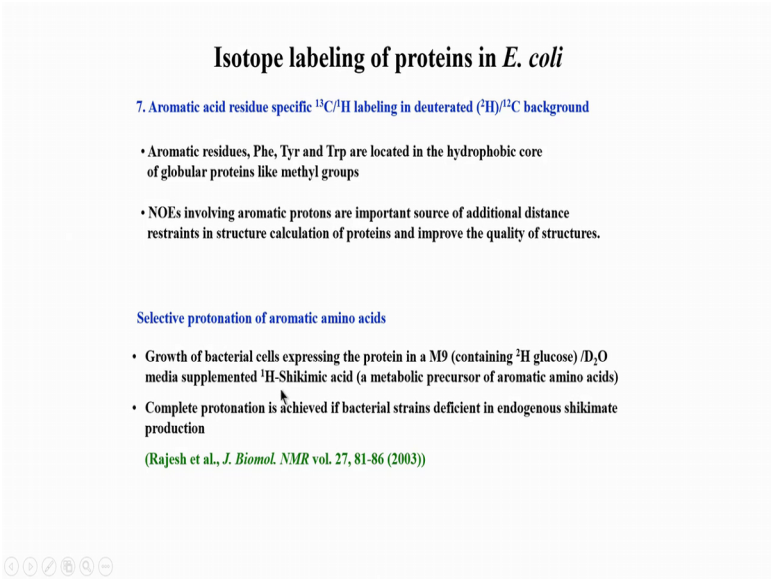
Because, there are no more hydrogens they will all become deuterons deuterium. So, we lose their information. So, what we do is we want to keep this in a protonated form; in that protonated form is kept and the remaining hydrogen's are all may deuterium. So, all the remaining hydrogens in the protein are deuterated and only the methyls are kept protonated. So, this is very useful for structure determination of proteins because, as shown here about 30 percent of the amino acids in a protein that is in a globular protein a structured protein typically are made up of methyl groups.

So therefore, methyl groups are very useful for structure determination and they are hydrophobic and almost always located in the interior of globular proteins and important

for protein folding. So now, if I get methyl to methyl this hydrogen to this hydrogen or this hydrogen to this and that is a NOESY that will help me to get distance information. So, how do we get information from NOESY, this is something we have seen in the 2 D part and we would again come back to it when actually we look at the 3 D part structure determination part of protein.

And the second reason why this CH methyl groups are very useful in proteins are because, they are very sensitive. Remember this each of these hydrogen there are 3 hydrogen is here. So, each hydrogen is 3 times stronger compared to remaining which are other a single hydrogen's. So, methyl groups are very strong sensitive and very easy to see them in the spectrum. Also they are well resolved because they are very sharp and therefore, it is always beneficial to keep the protons in the methyl group and deuterate the remaining part.

(Refer Slide Time: 17:11)



So, let us now continue with the other part which is basically how we can look at amino acids other than methyls. So, these are another set of amino acid called aromatic amino acids, they are also very useful in proteins. So, as shown here aromatic amino acid which are basically phenylalanine, tryptophan and tyrosine are located in the hydrophobic core. They are also hydrophobic amino acids similar to methyl group. So therefore, if I deuterate phenylalanine or tyrosine or tryptophan I will lose the hydrogens in this. And,

if I lose the hydrogens here I will not be able to get the distance information which is required for a structure determination.

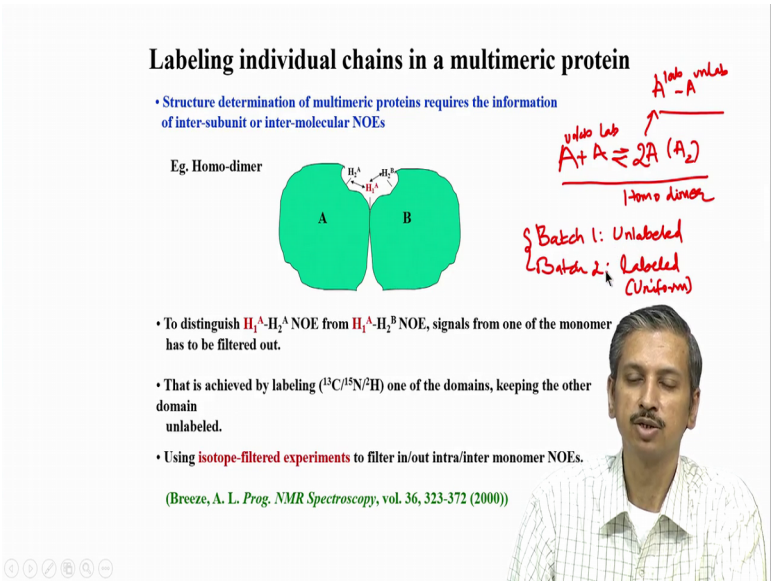
So, similar to methyls you want to keep this protonated. So, similarly here as mentioned here the NOEs involving the aromatic protons are important because, they give us additional distance information which is useful in structure calculation. We will see this in the structure part in the coming lectures and therefore, we need to have them protonated. So now, if you want to keep this protonated we have to get we should not deuterate them. So, how do we do that? We can show here in this protocol, we grow the bacterial cells in a M 9 that is in a minimal medium containing glucose that is deuterated glucose and we are using again remember D 2 O medium.

So, D 2 O medium is always present irrespective of what kind of labelling or deuteration you are doing. Now what do we do then, how do we get this in only in the protonated form? For that we use this particular amino acid sorry particular precursor, what is this precursor called shikimic acid. So, if we use a protonated shikimic acid means proton labelled means there is no deuteron here it is a metabolic precursor means it gives rise to amino acid. So, amino acids all come from shikimic acid pathway.

This is something which we you can actually see in that slide which earlier in previous lectures I had shown for the biosynthetic pathway and also is available in many biochemistry text book. So, I mean aromatic amino acids which come from the shikimic acid pathway you give supply shikimic acid in the medium and that gets converted to aromatic amino acid. But, now because it is protonated the aromatic acid now completely get protonated. So that means, you keep this hydrogen on the aromatic amino acids whereas, the remaining hydrogens in the protein are deuterated including methyls.

Because, remember here we are not doing methyl protonation we are doing aromatic acid protonation. So, methyls will get deuterated or all other hydrogens may get deuterated because, we are using deuterated glucose and D 2 O media. But, only the aromatic amino acids are remain they remain protonated because, of adding this particular metabolic precursor to the medium. So, this is how selective isotope deuteration is done for aromatic and methyls.

(Refer Slide Time: 20:25)



So, let us move on to the next isotope labelling scheme which is applicable very specifically to dimers. So, this is an example of a homo dimer. So, remember what are dimers? Dimers I showed you in the very beginning for protein structure in the second part of this course; when we started where we looked at their different structures of protein. One is called of course, one is primary secondary tertiary structure, but when it comes to quaternary structures also we have different categories. We had this IDPs unlabeled Intrinsically Disordered Proteins and we had this dimers and trimers and quaternary structures.

So, in dimer what happens is you have 2 proteins I mean the same protein homo dimer the same protein now exist in 2 form in in a dimer form. So, A and B are actually the same protein or it could be a hetero dimer hetero dimer could be A and B are different proteins. But, suppose let us say A and B are the same protein that is a homo dimer. Now, how do I get the structure of this complex, this is called a complex. So, for that I need to get the information of distance between the hydrogens from one protein to another protein this is very important in a dimer problem.

In dimers we need to know how the interface, interface means how this where the 2 molecules are interacting how that interface looks in 3 dimensions. So, for that I need to do specific labelling scheme to get that distance information between the 2 dimers. Remember I am interested between the dimers, I am not interested in that within one

monomer. Within one monomer anyway I can use standard labelling schemes, but now I want to get distance between the dimers how should I get. So, the problem here now is at there is a exact symmetry ok. So, when there is a symmetry in this molecule, in the dimer you cannot distinguish the signal of this side from the signal of this molecule.

Because, both are the same protein its only thing is they are dimer. So, that is a big problem in NMR is that the number of signals which come will be now 2 times one coming from A and another coming again from A. So, let me right this reaction. So, what is happening is 2 molecules of A are actually forming a dimer 2 A or A 2. So, this is a homo dimer because both are the same molecule. So, in a homo dimer the number of signals which you get are basically 2 times the signal of each of the molecules. Now, not only that this homo dimer if it has a symmetry axis which normal is present then the signals of this side cannot be separates or distinguish a signals from this side.

Both will have the same signal and therefore, it will double. So, how do I distinguish or how do I segregate or separate or discriminate between this protons and this protons because, if I cannot segregate then I will not be able to separate them. So, for that what we do we use this idea that you do grow the cells you grow one unlabeled protein that is completely unlabeled normal standard protein and second batch you make 2 batches. So, I can write this here you make batch 1, batch 1 unlabeled protein and batch 2 we make it labelled and labelled means uniform labelling it can be uniform or so on or it can also be site specific labelled.

So, we make 2 samples: one sample is unlabeled of this protein and second sample is labelled; now what we do is we mix the 2. So, when you mix the 2 then you will have one set of protein which are unlabeled and another molecule which are labelled, they will join together. And, now this will look like this labelled unlabeled pair. Why does this happen? Because, unlabeled from this batch if I mix with label of equal amount 50-50 or 1 is to 1 they will start mixing with each other randomly ok, because this dynamic this this dimer has to be is not see you can see this a equilibrium here.

So, it is not that it is sticking tightly they can be separated. So, and when you mix the 2 basically batch 1 and batch 2 which has; one is unlabeled set of samples another is labelled set. The unlabeled molecules will combine with labelled because everything can any A can combine with any A randomly and you will start seeing this pair in the sample

I mean this pair will be there in the sample. Of course, you will also have unlabeled unlabelled combination label labelled combination and so, on that is also possible, but this combination also will be there in the system.

In fact, this will be much more populated compared to label labelled or unlabeled combination. So therefore, that kind of a mixture is prepared ok. So, this is what is done here, this is done by labelling one of the domains or one of the one sample and keeping the other unlabeled. And, then we use the specific set of NMR experiment which we cannot go into detail, but this is just for your information. A specific set of NMR experiments are used which are then helpful and exploit this kind of a labelling. So, basically in this whole idea here is we are trying to get intermolecular NOE.

Why intramolecular, we want to get between the 2 molecules the distance information and that information is not available in a pure dimer form. So, I need to make a special sample in which I mix unlabeled batch 1 one batch or one set of sample prepared in unlabeled form with one which have prepared in the labelled form. And, when I mix these two I will get up I will end up with getting what is called a heterogeneous dimer. Because, this is no longer a homogeneous, they are not the same. They are two different molecules now because, one is labeled another is unlabeled.

So, by doing this I am able to now make this dimer exactly not same, they are now different. And, now I should be able to do some special experiments in NMR to distinguish these two because, the sample now helps me to do that. The sample helps me to separate the two and then the special experiments which we do are called isotope labelled experiment.

(Refer Slide Time: 27:21)

So, let us see what are those isotope labelled experiments; this are shown here. So, we basically how do we make the sample first? We mix equimolar ratio of the 2 monomers in any combination. So, I can take unlabeled monomer 1 I 1 monomer can be unlabeled or labelled, second monomer can be unlabeled; that means, second monomer means 2 batches the same monomer is mixed. Similarly I can have one suppose these two are not same I can take one in labelled, unlabeled or I can do the other one in unlabeled form take one in labelled form. Not only that you can see here I can mix them in different combinations, I can make N 15 labelled protein of one monomer, mix it with deuterated form of the second monomer.

Or I can take only C 13 labelled of 1 monomer and mix it with N 15 part of the second monomer. So, basically this side and this side are different batches means, I make a separate sample like this 1 monomer. I make a separate sample like this another monomer and I mix the two. Similarly, if you want to make the third combination here I make a separate sample, separate batch like this sample which is C 13 labelled. But, otherwise it is N 14 and I make a separate batch here and I mix the two. So, when I do that I will get a combination in which 1 monomer is this kind of labelling, second monomer is this. Now, with these kind of samples we can do this what is called isotope filter experiment and get the NOEs between the molecules. So, we are now getting inter and intra molecular NOE ok. So, we will not go into details of these two this labelling scheme.

But, they suffice here is to say that for dimers hetero dimers or homodimers we need to make different labelled samples of the monomer and mix them. And, once you mix them then you have special NMR experiments which can be used. I would recommend you to go through this article in this journal and that will give you much more details of how these experiments are performed on this different labelled samples ok. So now, this is basically the above samples need to be mixed; now what happens is sometimes these two cannot be separated easily. Because, they may be very tightly be bound to each other; if they are tightly bound to each other then you cannot separate and get a combination like this. Here I have assumed that they are easily breakable.

But that may not be the case every time. So, how do you solve in such cases? What you do is as shown here, we do the same thing like we did with deuteration we unfold the protein. Basically, what we do is we can unfold means we separate them by unfolding. How do you unfold a protein? We can add denaturants and what are denaturants; there are urea, guanidine hydrochloride and so on. So, we can denature this proteins; that means, these two will separate out and then I can denature one; the first batch of sample, I can denature the second batch separately then I can mix the two

So, when I after denaturing when I mix, I will end up with this combination. But, if I do not denature and then mix I would not be able because, in batch 1 unlabeled would have already formed a dimer with 1 labelled. Similarly in batch 2 a labelled would have already formed a dimer with labelled. So, again they will be homogeneous. To make it heterogeneous like this, I may have to denature this whole system that is unlabeled did unfold this, then mix the 2 batches and refold it. How do I refold it? I just remove that urea or guanidine hydrochloride by dialysis.

So, when I do dialysis and remove them the protein will start folding back, but now they will start combining and form heterogeneous mixtures like this. This is very specially done for those cases where, the dimers are very tight or high affinity. So, this brings us to the end of the bacterial labelling. We have looked at a few isotope labeling schemes, we are not covered all of them. In the next class we will look at how labelling is done in yeast and the cell free synthesis. Remember I said cell free synthesis is today very popular because, that saves lot of money and only specifically a particular protein can be made. We will look at that in the next class.